

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 02:11:42 ; Search time 1596.96 Seconds
(without alignments)
2307.302 Million cell updates/sec

Title: US-09-699-652-11
Perfect score: 273
Sequence: 1 ctctctctgcacgttcg.....tctgttgccaangcggt 273

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_estl:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_gss:*
- 13: em_gss_hum:*
- 14: em_gss_inv:*
- 15: em_gss_pln:*
- 16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147.4	54.0	618	10	BE490056
2	96.8	35.5	684	10	BE594921
3	96.8	35.5	746	10	BE594730
4	95.2	34.9	375	10	BE098681
5	87.4	32.0	643	9	AW651434
6	82	30.0	564	10	BE434955
7	77.2	28.3	544	10	BE432838
8	72.8	26.7	577	10	BE432502
9	64.8	23.7	830	10	BF254425
10	60.8	22.3	398	10	BE436266
11	58.8	21.5	213	10	T46320
12	55.2	20.2	737	12	BH537343
13	53.8	19.7	680	12	AQ690237
14	53.4	19.6	623	10	BM324154
15	49.4	18.1	557	10	BE593553
16	46.4	17.0	665	9	A1296265
17	46.4	17.0	717	10	BE443524

18	46.2	16.9	962	9	BE040022
19	45.4	16.6	645	9	BE293432
20	45.4	16.6	702	10	BF479168
21	44.8	16.4	729	12	AQ330709
22	42.8	15.7	768	10	BE912909
23	42.4	15.5	624	9	AI259731
24	41.4	15.2	492	9	AL380410
25	40.4	14.8	365	10	R15016
26	40.4	14.8	846	12	CNS010R1
27	39.8	14.6	420	10	BI995281
28	39.4	14.4	427	9	AV401151
29	39.4	14.4	720	10	BF494536
30	39.2	14.4	423	9	AU161359
31	39.2	14.4	445	9	AU166509
32	39.2	14.4	455	9	AU181431
33	39.2	14.4	513	12	AZ985567
34	39.2	14.4	711	9	AU102211
35	39.2	14.4	837	10	BI546997
36	39	14.3	458	9	AI665559
37	39	14.3	668	10	BI075779
38	38.8	14.2	395	10	BF189425
39	38.8	14.2	754	9	AU141282
40	38.8	14.2	768	10	BE584493
41	38.8	14.2	950	9	AL569763
42	38.6	14.1	559	12	BH416018
43	38.4	14.1	379	9	AA956343
44	38.4	14.1	448	9	AA997882
45	38.4	14.1	451	10	BI283395

ALIGNMENTS

RESULT 1

BE490056
LOCUS
DEFINITION
WHE0364_A03_A06Zs Wheat cold-stressed seedling cDNA library
Triticum aestivum cDNA clone WHE0364_A03_A06, mRNA sequence.

ACCESSION
BE490056
VERSION
BE490056.1
KEYWORDS
EST.
SOURCE
bread wheat.
ORGANISM
Triticum aestivum

REFERENCE
AUTHORS
Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L. and Tong,J.C.

TITLE
The structure and function of the expressed portion of the wheat genomes - Cold-stressed seedling cDNA library

JOURNAL
COMMENT
Unpublished (2000)
Contact: Olin Anderson

US Department of Agriculture, Agricultural Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595618

Email: oanderson@w.usda.gov
Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.

FEATURES
source
1..618
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE0364_A03_A06"
/clone_lib="Wheat cold-stressed seedling cDNA library"
/tissue_type="seedling"
/dev_stage="Five-day old seedling"
/lab_host="E. coli SOLR"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site_1: EcoRI; Site_2: XhoI; Seeds were surface-sterilized, germinated and grown aseptically in the dark at room temperature on filter paper with water, nystatin and cefotaxime in covered crystallization dishes. Five-day old seedlings were transferred to 5 C cold room and kept for 48 hr. The tissue, total RNA, and poly(A) RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript phagemids in the TJ Close lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 119 a 211 c 179 g 109 t
ORIGIN

Query Match 54.0%; Score 147.4; DB 10; Length 618;

Best Local Similarity 79.6%; Pred. No. 1.7e-21;
Matches 172; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 30 tactgaactggtctcgatgacctgtgtcaacgacctgcggccatggtcgacttc 89

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 111 TACTGGGACTGGTCTGGGAGCAGCTGGTGTGACGACATGCCGAGCATGTTGACTAT 170

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 90 gtgtcaaacagacggccagaagcctcactacgtcgacactcatggtgacgctggtg 149

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 171 ATCTCAGCCACACGGCGCAGAAAGCGCACATTCCTCGGCCACTCCATGGGACGCTGGTG 230

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 150 ggcgtgagcgcccttcgagagggcggtgtgtgagcagctgaaatccgcggcgctgctc 209

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 231 GCGCTGGCGGCGCTTCTCGAGGGGCAAGACGCTGGCAAGCTCAAGTCGGCGCGCTGCTG 290

Qy 210 acgcgggtggcctacctcgcacatnaacancccc 245

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 291 ACCCGGTGCGCTATCTCTCCCATGACACCCCC 326

RESULT 2

BG594921

LOCUS ESP493599 cSNS Solanum tuberosum cDNA clone cSTS874 5' sequence,
DEFINITION mRNA sequence.

ACCESSION

BG594921

VERSION BG594921.1 GI:13613061

KEYWORDS

EST.

SOURCE

potato.

ORGANISM

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

1 (bases 1 to 684)

van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemiango,A.,

Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.

Generations of ESTs from sprouting potato eyes

Unpublished (2000)

Contact: Cathy Ronning

The Institute for Genomic Research

For clone info: please contact Research Genetics, Libraries

Division tel 1-800-711-6195, email cdna@resgen.com

Seq primer: M13F-R.

Location/Qualifiers

1. .684

/organism="Solanum tuberosum"

/cultivar="Kennebec"

/db_xref="taxon:4113"

/clone="cSTS874"

/clone_lib="cSTS"

/tissue_type="sprouting eyes from tubers"

/dev_stage="12-14 weeks post harvest"

/lab_host="SOLR"

/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:

XhoI; Various sizes of sprouting eyes (2mm to 15mm) were

taken from tubers. The tubers were incubated at 26C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."

BASE COUNT 173 a 149 c 172 g 190 t
ORIGIN

Query Match

35.5%; Score 96.8; DB 10; Length 684;

Best Local Similarity

62.9%; Pred. No. 6.7e-11;

Matches 146; Conservative

0; Mismatches 86; Indels

0; Gaps

0;

Qy 29 ctactggaactggtctcggtgacgtgtgtgtagtcaacgacctgcggccatggtcgactt 88

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 208 CTACTGGAATTGGACATGGATGATCTTGTCTGCTCCATGACTTACCTGCTCTGTTGATCT 267

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 89 cgtgctcaaacagacggccagagccctcactacgtcgagacactccatcgaggacgctggt 148

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 268 TGTCTTTAAACAAACTGGGCAGAAATTCACATATATAGCCCATTCATATGGGAACCTTGAT 327

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 149 ggcgtgagcgccctctcgagggccggtgtgtgagccagctgaaatccgcggcgctgct 208

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 328 AGCGTGGCGTCTTTTCAGAAGGGAAACAAATAGACAAGGTAAATCAGACGCTTGCT 387

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 209 cacgcgggtggcctacctcgcacatnaacanccccacatnggaatcctggt 260

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 388 CAGCCCACTGCTTATTGAGCCATATGACCACCTGCTCGGTGATGTTGCT 439

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 3

BG594730

LOCUS EST493408 cSTS Solanum tuberosum cDNA clone cSTS8E14 5' sequence,

DEFINITION mRNA sequence.

ACCESSION

BG594730

VERSION BG594730.1 GI:13612870

KEYWORDS

EST.

SOURCE

potato.

ORGANISM

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

1 (bases 1 to 746)

van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemiango,A.,

Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.

Generations of ESTs from sprouting potato eyes

Unpublished (2000)

Contact: Cathy Ronning

The Institute for Genomic Research

For clone info: please contact Research Genetics, Libraries

Division tel 1-800-711-6195, email cdna@resgen.com

Seq primer: M13F-R.

Location/Qualifiers

1. .746

/organism="Solanum tuberosum"

/cultivar="Kennebec"

/db_xref="taxon:4113"

/clone="cSTS8E14"

/clone_lib="cSTS"

/tissue_type="sprouting eyes from tubers"

/dev_stage="12-14 weeks post harvest"

/lab_host="SOLR"

/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:

XhoI; Various sizes of sprouting eyes (2mm to 15mm) were

taken from tubers. The tubers were incubated at 26C in the

dark for 2-3 weeks prior to sprouting. The eyes were

frozen in liquid nitrogen immediately upon removal from

tubers."

BASE COUNT 198 a 163 c 169 g 216 t

ORIGIN

Query Match

35.5%; Score 96.8; DB 10; Length 746;

```

Best Local Similarity 62.9%; Pred. No. 6.8e-11;
Matches 146; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Qy 29 ctactggaactggtcctctggaatgacctgtagtcaacagacctgcccggccatggtcgactt 88
Dbb 73 CTACTGGAATTGGACATGGGATGATCTTGTGCTCATCATTCATTACCTGCTCTCGTTGATCT 132
Qy 89 cgtcgtcaaacagacccgcccagaagcctcactacgtcgagacactcatcctgggagcgtggt 148
Dbb 133 TGTCTTTAAACAACTGGCGCAGAAAATTCACATATATAGCCCATTCATGGAACCTTGAT 192
Qy 149 ggcgcgtggtcctctcggaggccgggtggtgagccagctgaaatccgcggcgctgct 208
Dbb 193 AGCGTGTGGCGTCTTTTCAGAGGAGGAAACAAATAGACAAGGTAAATCAGCAGCCTTGCT 252
Qy 209 cagccggtggtcctcctcgcgcacatnaacancncccaatnggaatcctggt 260
Dbb 253 CAGCCCACTGCTTATTGAGCCATATGACCACCTGCACCTCGGTGATGTTGCT 304

RESULT 4
LOCUS BG098681 375 bp mRNA linear EST 29-JAN-2001
DEFINITION BG098681 sprouting eyes/shoots Solanum tuberosum cDNA clone
cSTC5E1 5' sequence, mRNA sequence.
ACCESSION BG098681
VERSION BG098681.1 GI:12588716
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 375)
AUTHORS van der Hoeven,R.S., Bezzerides,J., Cho,J., Utterback,T., Hansen
,C.L., Bougri,O., Buell,C.R., Ronning,C.M., Tanksley,S.D. and Baker
,B.
TITLE Generation of ESTs from potato sprouting eyes/shoots
JOURNAL Unpublished (2001)
COMMENT Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
FEATURES
source
1. 375
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTC5E1"
/clone_lib="sprouting eyes/shoots"
/tissue_type="sprouting tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOUR"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Various sizes of sprouting eyes (2 mm to 15 mm) were
taken from tubers. The tubers were incubated at 26C in
the dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."
BASE COUNT 97 a 77 c 89 g 112 t
ORIGIN
Query Match 34.9%; Score 95.2; DB 10; Length 375;
Best Local Similarity 62.5%; Pred. No. 1.3e-10;
Matches 145; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy 29 ctactggaactggtcctctggaatgacctgtagtcaacagacctgcccggccatggtcgactt 88
Dbb 73 CTACTGGAATTGGACATGGGATGATCTTGTGCTCATCATTCATTACCTGCTCTCGTTGATCT 132
Qy 89 cgtcgtcaaacagacccgcccagaagcctcactacgtcgagacactcatcctgggagcgtggt 148

```

```

Dbb 133 TGTCTTTAAACAACTGGCGCAGAAAATTCACATATATATGCCATTCAATGGGAACCTTGAT 192
Qy 149 ggcgcgtggtcctctcggaggccgggtggtgagccagctgaaatccgcggcgctgct 208
Dbb 193 AGCGTGTGGCGTCTTTTCAGAGGAGGAAACAAATAGACAAGGTAAATCAGCAGCCTTGCT 252
Qy 209 cagccggtggtcctcctcgcgcacatnaacancncccaatnggaatcctggt 260
Dbb 253 CAGCCCACTGCTTATTGAGCCATATGACCACCTGCACCTCGGTGATGTTGCT 304

RESULT 5
LOCUS AW651434 643 bp mRNA linear EST 18-MAY-2001
DEFINITION EST329888 tomato germinating seedlings, TAMU Lycopersicon
esculentum cDNA clone cLE11617 5', mRNA sequence.
ACCESSION AW651434
VERSION AW651434.1 GI:7412672
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 643)
AUTHORS Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Niernan,W., Fraser,C.M., Giovannoni,J.J., Martin,G.B. and Tanksley
,S.D.
TITLE Generation of ESTs from germinating tomato seed
JOURNAL Unpublished (2000)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES
Location/Qualifiers
1. 643
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLE11617"
/clone_lib="tomato germinating seedlings, TAMU"
/tissue_type="whole seedlings"
/dev_stage="7 days post imbibition"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; 7 days post imbibition on water-agar. Mixed stage
whole germinating seedlings from seed coat emergence up
to two centimeters in seeds not showing obvious signs of
germination were discarded."
BASE COUNT 177 a 111 c 156 g 199 t
ORIGIN
Query Match 32.0%; Score 87.4; DB 9; Length 643;
Best Local Similarity 61.5%; Pred. No. 6.2e-09;
Matches 136; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Qy 30 tactggaactggtcctggaatgacctggtagtcaacagacctgcccggccatggtcgacttc 89
Dbb 422 TACTGGAATTGGTCATGGATGATTTGATTTGATGATTTTACCATCTGTTATTGACTTT 481
Qy 90 gtcgtcaaacagacccgcccagaagcctcactacgtcgagacactccatcgggagcgtggtg 149
Dbb 482 ATCTTCAACAACTGGCAGAGAAAATACACTATGTTGGTCATTCATCAATGGGAACCTTGATA 541
Qy 150 ggcgcgtggtcctcctcggaggccgggtggtgagccagctgaaatccgcggcgctgctc 209
Dbb 542 GCTTTGGCATCATCTCTCAGAGGAGGAAATAGACAAGGTAAATCAGCAGCATTTACTA 601
Qy 210 acgcgcgtggtcctcgtacatcgnccacatnaacancncccaatng 250

```

Db 602 AGTCCAATTGCTATTATGGCCATATGACCACTGCCTTGG 642

RESULT 6

BE434955

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

BE432838
EST399463 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG11C5, mRNA sequence.

Asteridae: euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.

REFERENCE
AUTHORS
Alcala, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E.,
Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Renning, C.M.,
Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley
S.D.

Development of a genetically and physically anchored EST resource
for barley genomics: Morex unstressed seedling root cDNA library
unpublished (2001)
On Nov 16, 2000 this sequence version replaced gi:13117126.
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 561
Seq primer: AATTAAACCTCTACTAAAGG
High quality sequence start: 564.

RESULT 9
 BF254425
 LOCUS
 DEFINITION
 HVSMEF0003P17f Hordeum vulgare seedling root EST library HVcDNA0007
 (Etisolated and unstressed) Hordeum vulgare cDNA clone
 HVSMEF0003P17f, mRNA sequence.
 ACCESSION
 VERSION BF254425
 KEYWORDS BF254425.3 GI:16315084
 SOURCE EST.
 ORGANISM
 Hordeum vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
 1 Triticeae; Hordeum.
 1 (bases 1 to 830)
 REFERENCE
 AUTHORS Wang,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu,
 Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W., Fenton
 R.D., Oates,R. and Main,D.

	Query Match	23.7%	Score 64.8;	DB 10;	Length 830;
	Best Local Similarity	56.7%;	Pred. No. 0.00035;		
	Matches 136;	Conservative	0;	Mismatches 101;	Indels 3; Gaps 1;
QY	2	tctctentgcacggttcgntttcagctctactgaacctgtccctggatgacctggtagt	61		
Db	395	TACCTCGCTCACTCCAGATGATCCGGAGTACTGGAACTGGACGTGGGACCNACTTGCTGC	454		
QY	62	caagacactcgcgccaatggctgactctgctgtctaaacagacc---	ggccagaagacctca	118	
Db	455	CTATGATCTTCCTGCCGTGCTCAGTTTGTTCTATGATCACACAGGAGGCCAAGAATCCA	514		
QY	119	ctacgtcgacatcccatgggacgctggggcgctgcgcgcgcctctctcgagggcccggt	178		
Db	515	CTATATCGGTCATCCCTCGGGAACTTGATTATTATTTGCAGCCCTCTCTGACACACAGGTT	574		
QY	179	ggtgagccagtgaataccgcgcgctgcctcacgcggtggcctacctcgcccacatnaa	238		
Db	575	ACTCATTTAGTTCGATCGGTTGGTGGCTCTGCCAAATGCTTATCTGTATAAGACAAA	634		

```

RESULT 10
BE436266
LOCUS BE436266 398 bp mRNA linear EST 18-MAY-2001
DEFINITION EST407344 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
ACCESSION BE436266
VERSION BE436266
KEYWORDS EST.
SOURCE EST.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 398)
Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Niernan,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
,S.D.
Generation of ESTs from tomato fruit tissue, breaker stage
Unpublished (2000)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES
source
1..398
Location/Qualifiers
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLEG31C20"
/clone_lib="tomato breaker fruit, TIGR"
/tissue_type="pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/note="Vector: pBluescriptSKmUadapT; Site.1: EcoRI;
Site.2: XhoI; Fruit were harvested at the breaker stage
(first sign of lycopene accumulation on the blossom end of
the fruit). Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."
BASE COUNT 112 a 62 c 103 g 121 t
ORIGIN
Query Match 22.3%; Score 60.8; DB 10; Length 398;
Best Local Similarity 67.2%; Pred. No. 0.0021;
Matches 86; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
Qy 30 tacttgaactggtcctggatgacctgtagtcaacgacctgcggccatggtcgacttc 89
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 268 TACTGGAATTGTCATGGGATGAATTGATTGTCATGATTACCATCTGTTATTGACTTT 327
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 90 gtcgtcaaacagaccggccagaagcctcactacgtcgacacctcatcggaacctcgtgtg 149
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 328 ATCTTTAAACAACCTGGACAGAAATACACATATGTTGTCATTCAATGGAACTATGATA 387
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 150 gcgcctggc 157
| | | | |
Db 388 GCTTTGGC 395

RESULT 11
T46320
LOCUS T46320 213 bp mRNA linear EST 04-AUG-1998
DEFINITION 9583 Lambda-PRL2 Arabidopsis thaliana cDNA clone 141F5T7, mRNA
sequence.
ACCESSION T46320
VERSION T46320.1 GI:2763011
KEYWORDS EST.
SOURCE thale cress.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 213)
Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel
,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729
On Jan 9, 1998 this sequence version replaced gi:934542.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-9168
Fax: 517-353-0854
Email: 22313tcn@bm.cl.msu.edu
Seq primer: T7 dye primer.
FEATURES
source
1..213
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone="141F5T7"
/clone_lib="Lambda-PRL2"
/note="Vector: lambda zip-Lox; Site.1: Sal; Site.2: Not;
Lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques). The vector is BRU's lambda zip-Lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dT primed cDNA."
BASE COUNT 49 a 53 c 50 g 53 t 8 others
ORIGIN
Query Match 21.5%; Score 58.8; DB 10; Length 213;
Best Local Similarity 58.8%; Pred. No. 0.0051;
Matches 90; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
Qy 102 accggccagaagcctcactacgtcggaacctcatcgggagcgtggtggcgcgcc 161
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 18 ANAGCCAAAANATTCACTCGTNACTCTTTGGGACGTTAATAGGATTTCTCTCG 77
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 162 ttctcgaggaggcggtgtgtgagccagctgaaatccgcgagctgcacccgctgcacccggtgcc 221
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 78 TTCTCAGAAAAGGGTTGGTGATCAAGTNAGATCGCGGGATGTTGATCCCCGTGCT 137
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 222 tacctcgccacatnaacanccccattngaat 254
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 138 TATCTNAGCCCATGACCCACCGTNATCGCGAT 170
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 12
BH537343
LOCUS BH537343 737 bp DNA linear GSS 14-DEC-2001
DEFINITION BOGCC23TR BOGC Brassica oleracea genomic clone BOGCC23, DNA
sequence.
ACCESSION BH537343
VERSION BH537343.1 GI:17773756
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.

```

REFERENCE	1 (bases 1 to 737)
AUTHORS	Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
TITLE	Whole genome shotgun sequencing of <i>Brassica oleracea</i>
JOURNAL	Unpublished (2001)
COMMENT	Other_GSSS: BOGCC23TF Contact: Chris Town

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtownet@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

```

FEATURES             Location/Qualifiers
     source            1..737
                        /organism="Brassica oleracea"
                        /strain="TO1000DH3"
                        /db_xref="taxon:3712"
                        /clone="BOGCC23"
                        /clone_lib="BOGC"
                        /note="vector: PHOS1; Site_1: BstXI; 2-3 kb
genomic DNA inserted into PHOS1 using BstXI
197 a 154 c 154 g 232 t

```

BASE COUNT	197 a	154 c	154 g	232 t	genomic DNA inserted into phage using BSA1 linkers
ORIGIN					

[illegible]

ACCESSION	AQ690237
VERSION	AQ690237.1
KEYWORDS	GI:5331405
SOURCE	GSS.
ORGANISM	Oryza sativa.
	Oryza sativa.
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
AUTHORS	1 (bases 1 to 680)
TITLE	Wing, R. A. and Dean, R. A.
JOURNAL	A BAC End Sequencing Framework to Sequence the Rice Genome
COMMENT	Unpublished (1998)
	Contact: Wing RA

FEATURES
source 1. .680

```
/organism="Oryza sativa"  
/strain="Japonica"  
/cultivar="Nipponbare"  
/db_xref="taxon:4530"
```

```

/clone="nbxb0081N07f"
/clone_lib="CUGI Rice BAC Library"
/tissue_type="leaf"
/lab_host="E. coli DH10B"
/notes="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
HindIII; Rice is one of two most popular grains in the
world. Half of the world population especially those
inhabiting highly populated areas of the humid tropics
and subtropics, rely on rice as their primary source of
carbohydrate. Monocotyledonous rice is a diploid plant
(2n=24) with a haploid genome equivalent of 431 Mbp
(Arumuganathan and Earle, 1991). The relatively small
genome of rice, three times larger than that of
Arabidopsis, makes it suitable for genomic studies. In
order to facilitate positional cloning, physical mapping
and genome sequencing of rice, we have constructed a BAC
library from Oryza sativa, Nipponbare variety. The
library contains 36,864 clones with an average insert size
of 128.5 Kb providing 10.9 haploid genome equivalents. The
deep coverage allows the isolation a particular sequence
with a probability of 99.9 %. Two high density filters,
each containing 18,432 clones (doubly spotted), represent
the whole library for colony screening."
```

BASE COUNT	211 a	143 c	154 g	172 t
ORIGIN				

	Query Match	19.7%	Score 53.8	DB 12	Length 680
	Best Local Similarity	65.3%	Pred. NO. 0.068		
	Matches	79	Conservative	0	Mismatches 42
				Indels	Gaps
					0
Qy	21	tttcagctctactggaactggtctctgggagacctggtagtcaacgacctgcgcggccatg	80		
Db	325	TTTCAGGCTTACTGGGAATGGTCGTGGGAGCACTGCTTCCNTACGATCTCTGCAGTG	266		
Qy	81	gtcgacttcgtctcaaacagacgcggccagaagcctcactacgtctcgagacatccatgggg	140		
Db	265	CTGCAGTTTGCCATGATCACACAGGGCGAGAAATCCACTATATCGTCACTCCCTGGTG	206		
Qy	141	a 141			
Db	205	A 205			

RESULT	14
BM324154	
LOCUS	BM324154
DEFINITION	PIC1_24_G11.bl_A002 Pathogen-infected compatible 1 (PIC1) Sorghum bicolor cDNA, mRNA sequence.
	623 bp
	mRNA
	linear
	EST 04-JAN-2002

ACCESSION	BM324154
VERSION	BM324154.1
KEYWORDS	GI:18062592
SOURCE	EST.
	sorghum.

ORGANISM

Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE

1 (bases 1 to 623)

AUTHORS	Cordonnier-Pratt, M.-M., Gingle, A., Fang, G.C., Dean, R., Wing, R., Sudman, M. and Pratt, L.H.
TITLE	An EST database from Sorghum: plants infected with a compatible pathogen
JOURNAL	Unpublished (2002)
COMMENT	Contact: Cordonnier-Pratt MM

Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu

T7 sequencing primer, are presented as the reverse complement.

Seq primer: JEN REV
High quality sequence stop: 573
POLYA-No.

FEATURES
source
Location/Qualifiers
1..623
/organism="Sorghum bicolor"
/cultivar="BTx623"
/db_xref="taxon:4558"
/clone_lib="Pathogen-infected compatible 1 (PIC1)"
/tissue_type="Leaves"
/dev_stage="4-week-old seedlings infected with Colletotrichum graminicola"
/note="Vector: pBluescript II SK(-) from Lambda Zap II; Site 1: XhoI; Site 2: EcoRI; Four-week-old sorghum seedlings were sprayed with spore suspension prepared from 3-week-old FRM421, a sorghum isolate of the anthracnose pathogen Colletotrichum graminicola. Inoculated plants were kept in a 25 C dark growth chamber with 100% relative humidity for 24 hr, followed by 12/12 hr of light/dark cycle at 25 C with 90% relative humidity for another 24 hr. All leaves were harvested and quick frozen with liquid nitrogen and stored in a -80 C freezer. The library was made from poly-A RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision. WARNING: While most or all ESTs are expected to derive from the host plant, no effort was made to eliminate ESTs deriving from the pathogen."
BASE COUNT 158 a 116 c 139 g 210 t
ORIGIN

Query Match 19.6%; Score 53.4; DB 10; Length 623;
Best Local Similarity 51.1%; Pred. No. 0.082;
Matches 120; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

Qy 3 tctctntgcagcttcgntttcagctctactggaactgctcctggatgacctgtagtc 62

Db 180 TCTACTCTCTCTGTCATGATAAGCTTTTCTGGGAATGGAGTTGGCAAGAGCCTTGTGAA 239

Qy 63 aacgacctgcgcgcagctgctgcacttcgctcaacagacgcggccagaagcctcactac 122

Db 240 TACGATGTTTGGCAATGTTAAGCTATGTATATACAAATACACATGCCAAATTCATAT 299

Qy 123 gtccgacactccatgggacgcgtggtgcgctgcggcctctcgcgagggccgggtggtg 182

Db 300 GTGGGACATTCACAGGAACATCATGCTGGTCTGGCTGTACAAATGCCCTGAATAGTA 359

Qy 183 agccagctgaaatccgcgcgcgtgctcgcgcgcgggtgacctcgcgcacatna 237

Db 360 AAAATGATAAGCTCGCTGTGCTCTTTGTGCCATTTCTTACCTTGATCATCA 414

RESULT 15

BG593553

LOCUS BG593553 557 bp mRNA linear EST 12-APR-2001
DEFINITION EST492231 cSTS Solanum tuberosum cDNA clone CSTS4D6 5' sequence,
mRNA sequence.

ACCESSION BG593553

VERSION BG593553.1 GI:13611693

KEYWORDS EST.

SOURCE potato.

ORGANISM Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

REFERENCE

AUTHORS

1 (bases 1 to 557)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemiango,A.,

Bougril,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.

Generations of ESTs from sprouting potato eyes

Unpublished (2000)

Contact: Cathy Ronning

The Institute for Genomic Research

For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.

FEATURES
source
Location/Qualifiers
1..557
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="CSTS4D6"
/clone_lib="cSTS"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."

BASE COUNT 147 a 115 c 145 g 150 t
ORIGIN

Query Match 18.1%; Score 49.4; DB 10; Length 557;

Best Local Similarity 71.2%; Pred. No. 0.55;

Matches 79; Conservative 0; Mismatches 31; Indels 1; Gaps 1;

Qy 29 ctactggaactgctcctggatgacctgtagtcaacgacctgcggccatgtgcactt 88

Db 447 CTACTGGGAATTGGACATGGATGATCTTGTGTCGCATGACTTACCTCTCGTTGATCT 506

Qy 89 cgtcgtcaaacagacgcggccagaagcctcactacgtcggacactccatggg 139

Db 507 TGTCCTTTAAACA-CTGGGCAGAAAAATTCACTATATAGGCCATTCAATGGG 556

Search completed: October 10, 2002, 02:41:26

Job time: 1784 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 02:13:42 ; Search time 38.97 Seconds
(without alignments)
1720.758 Million cell updates/sec

Title: US-09-699-652-11

Perfect score: 273
Sequence: 1 cttctctntgcaagcgttcgn.....tctcggttgccaangcgtt 273

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000
Maximum DB seq length: 200000000[illegible]

Post-processing: Minimum Match 08

Maximum Match 100%

1. **Introduction**

Database : Issued_Patents_NA:*

```
1: /cgn2_6/ptodata/2/1na/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B_COMB.seq:*
```

```
2: /cgn2_6/ptodata/2/ina/3B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
```

```
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
```

5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.sec

6: /cgn2_6/ptodata/2/ina/backfiles1.se

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	Time				
1	40.6	14.9	2987	4	US-09-105-537-40		Sequence 40, Appl
2	40.6	14.9	5970	3	US-09-320-878-21		Sequence 21, Appl
3	37.6	13.8	1578	1	US-08-681-129-1		Sequence 1, Appl
4	37.6	13.8	2085	1	US-08-553-999B-1		Sequence 1, Appl
5	37	13.6	30001	1	US-08-125-468-1		Sequence 1, Appl
6	37	13.6	30001	2	US-08-474-933-1		Sequence 1, Appl
7	36	13.2	4004	3	US-08-888-077A-30		Sequence 30, Appl
8	35.6	13.0	3283	4	US-09-651-656-16		Sequence 16, Appl
9	35.6	13.0	15239	1	US-08-390-878-17		Sequence 17, Appl
10	35.4	13.0	501	1	US-08-163-181-3		Sequence 3, Appl
11	35.4	13.0	501	1	US-08-465-161-3		Sequence 3, Appl
12	35.4	13.0	552	1	US-08-163-181-4		Sequence 4, Appl
13	35.4	13.0	552	1	US-08-465-161-4		Sequence 4, Appl
c 14	35.2	12.9	1977	3	US-08-371-529-2		Sequence 2, Appl
c 15	35.2	12.9	1977	4	US-08-977-816-2		Sequence 2, Appl
16	35	12.8	4403765	4	US-09-103-840A-2		Sequence 2, Appl
17	34.8	12.7	1174	3	US-09-034-985-1		Sequence 1, Appl
18	34.8	12.7	4411529	4	US-09-103-840A-1		Sequence 1, Appl
c 19	34.6	12.7	651	4	US-08-998-416-883		Sequence 883, App
20	34.6	12.7	1357	2	US-08-447-430A-36		Sequence 36, Appl
21	34.4	12.6	1137	2	US-08-227-108-2		Sequence 2, Appl
22	34.4	12.6	1137	2	US-09-073-674-2		Sequence 2, Appl
23	34.4	12.6	1140	1	US-08-227-108-4		Sequence 4, Appl
24	34.4	12.6	1140	2	US-09-073-674-4		Sequence 4, Appl
25	34.4	12.6	1146	1	US-08-227-108-6		Sequence 6, Appl
26	34.4	12.6	1146	2	US-09-073-674-6		Sequence 6, Appl
27	34.4	12.6	1528	1	US-08-227-108-1		Sequence 1, Appl

28	34.4	12.6	1528	2	US-09-073-67A-1	Sequence 1, Appl
C 29	34.4	12.6	2476	1	US-08-749-862A-1	Sequence 1, Appl
C 30	34.4	12.6	2476	1	US-08-539-134-1	Sequence 1, Appl
C 31	34.4	12.6	2476	5	PCT-US95-06816-1	Sequence 1, Appl
C 32	34.2	12.5	1357	2	US-08-447-430A-34	Sequence 34, Appl
C 33	34.2	12.5	43280	2	US-08-804-227C-1	Sequence 1, Appl
C 34	34	12.5	23673	4	US-08-773-816-1	Sequence 1, Appl
C 35	33.8	12.4	1833	3	US-08-403-852D-6	Sequence 6, Appl
C 36	33.8	12.4	1833	3	US-08-510-646B-6	Sequence 6, Appl
C 37	33.8	12.4	1833	4	US-09-231-818-6	Sequence 6, Appl
C 38	33.8	12.4	3003	4	US-08-915-337-1	Sequence 1, Appl
C 39	33.8	12.4	5247	1	US-08-920-812-15	Sequence 15, Appl
C 40	33.8	12.4	5247	1	US-08-920-827-15	Sequence 15, Appl
C 41	33.8	12.4	5247	1	US-08-921-177-15	Sequence 15, Appl
C 42	33.8	12.4	5247	1	US-08-362-577C-15	Sequence 15, Appl
C 43	33.8	12.4	5247	2	US-08-920-828-15	Sequence 15, Appl
C 44	33.8	12.4	4403765	4	US-09-103-840A-2	Sequence 2, Appl
C 45	33.8	12.4	4411529	4	US-09-103-840A-1	Sequence 1, Appl

ALIGNMENTS

```

RESULT      1
US-09-105-537-40
: Sequence 40, Application US/09105537A
: Patent No. 6265202
: GENERAL INFORMATION:
: APPLICANT: Sherman, D.H.
: APPLICANT: Liu, H.
: APPLICANT: Xue, Y.
: APPLICANT: Zhao, L.
: TITLE OF INVENTION: DNA encoding methymycin and pikromycin
: FILE REFERENCE: 600,438US1
: CURRENT APPLICATION NUMBER: US/09/105,537A
: CURRENT FILING DATE: 1998-06-26
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 40
: LENGTH: 2787
: TYPE: DNA
: ORGANISM: Streptomyces venezuelae
US-09-105-537-40

```

```

Query Match      14.9%; Score 40.6; DB 4; Length 2787;
Best Local Similarity 51.1%; Pred. No. 0.057;
Matches 94; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

Qy  49  atgacctggtagtcaacagaccctgcggccatggtcgacttcgtctgaacagaccggcc 108
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  1592 aggccagggcgggtcttcacggggggcccggaacgctcgctgcgagccgaac 1651
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  109 agaagcctcactagtcggagacactccatgggagcgcgtgtggcctgcggcccttcgg 168
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  1652 agatcctgcagggtgccggctgtcggaggagactcagagccctcgagagcgccctct 1711
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  169 agggccggggtggtgagccagctgaaatccgcgcgcgtactcagccgggtggcctacctcg 228
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  1712 tgggtctcgtccacgcgcgaccggtctcgaccggcgctgttctggtcgagccctgctcg 1771
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy  229 ncca 232
      | |
Db  1772 ccga 1775

RESULT 2
US-09-320-878-21
; Sequence 21, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.

```

; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320,878A
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 5970
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-320-878-21

Query Match 14.9%; Score 40.6; DB 3; Length 5970;
Best Local Similarity 51.1%; Pred. No. 0.066;
Matches 94; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

Qy 49 atgacctggtagtcacagacctcgccgcatggtcgcacttctgctcgtcgtcaaacagacccggcc 108
Db 4330 agggccagggcgggcgctctctccagcggggcccgaacagcctgctcgcgagggccgaac 4389

Qy 109 agaacctcactaactcgcggacactccatgggagacgtggtggcgctggtggcctctcgg 168
Db 4390 agatcctgcaggcgctcgcgctgtcggagagacgtacagagccctggagacggccctct 4449

Qy 169 agggccgggtggtgagccagctgaaatccggcgcgctgctcagccggtggtgacctcgcg 228
Db 4450 tggctcctcgtccacccgacccggctcgcgacccggcgctgtctggtcgcgagccctcgcg 4509

Qy 229 ncca 232
Db 4510 ccga 4513

RESULT 3
US-08-681-129-1
; Sequence 1, Application US/08681129
; Patent No. 5738854
; GENERAL INFORMATION:
; APPLICANT: Mettenleiter, Thomas Cristoph
; TITLE OF INVENTION: Pseudorabies virus vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Organon Teknika Corporation
; STREET: 1330 Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/681,129
; FILING DATE:
; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/244,446
; FILING DATE: 02-JUN-1994
; CLASSIFICATION: 424
; APPLICATION NUMBER: EP 92.203.079.6
; FILING DATE: 06-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 258-5200
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1578 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Pseudorabies virus
; STRAIN: Kaplan
; US-08-681-129-1

Query Match 13.8%; Score 37.6; DB 1; Length 1578;
Best Local Similarity 51.5%; Pred. No. 0.3;
Matches 85; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Qy 69 ctgcccggccatggtcgacttgctcaaacagacccggccagaagcctcactacgtcgga 128
Db 283 CTGACGGAGCTGGACCTCTGCGGCGCGGAGCTGTAGGACCCCTACCTCGTC 342

Qy 129 cactccatgggagacgtggtggcgctggcgccctctcggagggcggtggtgagccag 188
Db 343 GAGTCATGCTCTGCTGCGGCGTCTGCTGACGCTCTGTCACACGCCGCCAG 402

Qy 189 ctgaatccgcggcgctgctcagccggtggcctacccctcgnccac 233
Db 403 CAGGACCGCGTCTGGAGCTCTGCGCGGAGCCGCCCATCTGTCAC 447

RESULT 4
US-08-553-999B-1
; Sequence 1, Application US/08553999B
; Patent No. 5773599
; GENERAL INFORMATION:
; APPLICANT: Govindaswamy Chinnadurai
; TITLE OF INVENTION: Isolation and characterization of cDNA for E1A C-terminal b
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: G. Chinnadurai
; STREET: 3681 Park Avenue
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch; 713 Kb Storage
; COMPUTER: IBM Compatible (Gateway 2000)
; OPERATING SYSTEM: Windows 3.1
; SOFTWARE: Word for Windows (6.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/553,999B
; FILING DATE: 11/06/95
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5773599e
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2085 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: human
; INDIVIDUAL ISOLATE: PACT30
; CELL TYPE: B-cell
; IMMEDIATE SOURCE:
; LIBRARY: CDNA
; CLONE: PACT30
; PUBLICATION INFORMATION:
; AUTHORS: Ute Schaeper
; AUTHORS: Janice M. Boyd
; AUTHORS: Sulekha Verma
; AUTHORS: Erik Uhlmann
; AUTHORS: T. Subramanian
; AUTHORS: G. Chinnadurai
; TITLE: Molecular cloning and characterization of a cellular phosphoprotein that
; JOURNAL: Proceedings of the National Academy of Sciences United States of America
; VOLUME: 92
; ISSUE: 23
; PAGES: 10667-10671
; RELEVANT RESIDUES IN SEQ ID NO: 1:
US-08-553-999B-1

Query Match 13.8%; Score 37.6; DB 1; Length 2085;
Best Local Similarity 48.1%; Pred. No. 0.32;
Matches 100; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

Qy 51 gacctgtagtcaagacacgtgcccacatggtcgacttcgtcgaacagacggccag 110
Db 755 GGCCTCAACGAGACACACACACCTCATCAACGACCTCACCGCTCAACGACGATGAGACAA 814
Qy 111 aagcctcactcactcggaacacactcctgggacgctggtggcgctggcgcccttcctggag 170
Db 815 GGGCCCTCTCGTGTGAACACAGCCGGGTGCGCTGTGGATGAGAGGGCGCTGGCCCCAG 874
Qy 171 ggcgggtgtagcagcgtgaaatccgcgctgctcagccggtggtgacctcgcgc 230
Db 875 GCCCTGAAGGAGGCGCGATCCGCGCGCGCCCTGGATGTGCACGAGTGCGAACCCCTC 934
Qy 231 cacatnaacanccccgaatnagaatcctg 258
Db 935 AGCTTTAGCCAGGCCCTCTGAGGATG 962

RESULT 5
US-08-125-468-1
; Sequence 1, Application US/08125468
; Patent No. 5589385
; GENERAL INFORMATION:
; APPLICANT: Ryan, Michael J.
; APPLICANT: Lotvin, Jason A.
; APPLICANT: Strathy, Nancy
; APPLICANT: Fantini, Susan E.
; TITLE OF INVENTION: Cloning of the biosynthetic pathway for
; TITLE OF INVENTION: chlortetracycline and tetracycline formation and cosmids
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07470
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/125,468
; FILING DATE: 22-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsevdos, Estelle J
; REGISTRATION NUMBER: 31,145
; REFERENCE/DOCKET NUMBER: 31,255-02
; TELEPHONE: (201)831-3241
; TELEFAX: (201)831-3305
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-125-468-1

Query Match 13.6%; Score 37; DB 1; Length 30001;
Best Local Similarity 54.9%; Pred. No. 0.77;
Matches 73; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 60 gtaacgacacgtcgcccatggtcgacttcgtcgaacagacggccagacccctcac 119
Db 2355 GTCAATGACCGGGTCCGCGAGCGTCCGACCGCGCGGCGGACCCAGACCGACTCCGACGAATC 2314
Qy 120 tacgtcgacacactcatgggacgctggtggcgctggcgcccttcctcgagggcggtg 179
Db 2315 GAGCGCCAGCGCGCATCACCATCCGCTCCGCGGTGCTCTTCAACCGTGGCGGACGTC 2374
Qy 180 gtgagccagctga 192
Db 2375 AAGGTCAACCTCA 2387

RESULT 6
US-08-474-933-1
; Sequence 1, Application US/08474933
; Patent No. 5866410
; GENERAL INFORMATION:
; APPLICANT: Ryan, Michael J.
; APPLICANT: Lotvin, Jason A.
; APPLICANT: Strathy, Nancy
; APPLICANT: Fantini, Susan E.
; TITLE OF INVENTION: Cloning of the biosynthetic pathway for
; TITLE OF INVENTION: chlortetracycline and tetracycline formation and cosmids
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07470
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,933
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/125,468
; FILING DATE: 22-SEP-1993
; ATTORNEY/AGENT INFORMATION:

Query Match 13.2%; Score 36; DB 3; Length 404;
Best Local Similarity 52.7%; Pred No. 0.59;
Matches 69; Conservative 1; Mismatches 61; Indels 0; Gaps 0;

RESULT 8
US-09-651-656-16
: Sequence 16, Application US/09651656
: Patent No. 6340566

```

RESULT      8
US-09-651-656-16
; Sequence 16, Application US/09651656
; Patent No. 6340566
; GENERAL INFORMATION:
; APPLICANT: MCCUTHEN-MALONEY, SANDRA
; APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY
; TITLE OF INVENTION: DETECTION AND QUANTITATION OF SINGLE NUCLEOTIDE
; TITLE OF INVENTION: POLYMORPHISMS, DNA SEQUENCE VARIATIONS, DNA MUTATIONS,
; TITLE OF INVENTION: DNA DAMAGE AND DNA MISMATCHES
; FILE REFERENCE: IL-10689
; CURRENT APPLICATION NUMBER: US/09/651,656
; CURRENT FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/192,764
; PRIOR FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 3283
; TYPE: DNA
; ORGANISM: Thermus thermophilus
; US-09-651-656-16

```

```

Query Match      13.0%; Score 35.6; DB 4; Length 3283;
Best Local Similarity 49.7%; Pred. No. 1.1;
Matches 89; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 53 cctggtagtcaacgacctgcccggcatggtcgacttgcgtcgtcaaacagacccgccagaa 112
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 871 catgcccgggataccccataaggcgctttgacgcctacgcgaagcgtctctaagatggg 930
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 113 gccctcactacgtcggacactcactcattggggacgctggtgctgcgtgcgctctctcggaggg 172
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 931 ctccgcgctcggtggccgaccaggtggagcctgcgcgaggaggcggagcctcgtgtcg 990
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 173 ccggggtggtgagccagctgaaatccgcggcgctgctctcacgcggtggcctacctgcgcc 231
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 991 tcgggaggtgaccacagctcctcaccccggggaacctcacccaggaggscctctctcccc 1049
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
US-08-390-878-17
; Sequence 17, Application US/08390878
; Patent No. 5700683
; GENERAL INFORMATION:

```

REFERENCES

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 02:14:32 ; Search time 179.61 Seconds
(without alignments)
2609.641 Million cell updates/sec

Title: US-09-699-652-11
Perfect score: 273
Sequence: 1 cttctctgtgcacgttcgtn.....tctgtgtggccaangcgtt 273

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	266	97.4	273	21	AAZ34955
2	98.2	36.0	1718	21	AAZ34958
c	57	20.9	2608	23	ABL24956
3	52.8	19.3	1483	21	AAZ34956
5	52	19.0	705	21	AAZ34952
6	51.4	18.8	1651	23	ABLI7801
c	51.4	18.8	6656	23	ABLI7800
8	50.2	18.4	859	21	AAZ34953
9	49.4	18.1	25360	22	AAF88314

10	49.4	18.1	29736	22	AAF88317	S. spinosa DNA fra
11	48.8	17.9	1371	23	ABL24281	Drosophila melanog
12	48.8	17.9	5137	23	ABL24280	Drosophila melanog
13	40.6	14.9	2787	21	AAZ87302	S. venezuelae macr
14	40.6	14.9	5970	21	AAZ75635	Nucleotide sequenc
15	40.6	14.9	5970	21	AAZ56003	Contig 002 from co
16	40.4	14.8	2795	22	AAI71296	Streptomyces grise
17	40.4	14.8	2795	22	AAI71296	Streptomyces grise
18	40.4	14.8	5054	22	AAI81954	Pseudomonas putida
c	40.4	14.8	6453	22	AAI81954	Pseudomonas oleovo
19	40.4	14.8	6453	22	AAI81954	Corn sterol acyltr
20	40	14.7	311	22	AAI81954	Corn putative leci
21	40	14.7	921	21	AAI81954	Corn putative leci
22	40	14.7	1155	22	AAI81954	Corn putative leci
23	40	14.7	1660	21	AAI81954	Human cDNA sequenc
24	39.4	14.4	2651	22	AAI81954	Streptomyces globi
25	38.8	14.2	42000	21	AAI81954	Streptomyces globi
26	38.8	14.2	63164	21	AAI81954	Streptomyces globi
27	38.6	14.1	4446	21	AAI81954	S. aggregatum PKS
28	38.6	14.1	53789	19	AAI81954	Soybean acid triac
29	38.2	14.0	1438	21	AAI81954	Anycolatopsis medi
30	37.8	13.8	1194	22	AAI81954	Pimaricin biosynth
31	37.8	13.8	1374	23	AAI81954	Drosophila melanog
32	37.8	13.8	1521	19	AAI81954	DNA encoding trypt
33	37.8	13.8	3432	23	AAI81954	DNA encoding novel
c	37.8	13.8	3497	23	AAI81954	Drosophila melanog
35	37.6	13.8	1092	22	AAI81954	Streptomyces sp. C
c	37.6	13.8	1104	22	AAI81954	Streptomyces sp. C
36	37.6	13.8	1578	15	AAI81954	Nucleotide sequenc
37	37.6	13.8	1578	15	AAI81954	Open reading frame
38	37.6	13.8	2085	19	AAI81954	Nucleotide sequenc
39	37.6	13.8	2198	24	AAI81954	cDNA sequence #141
c	37.6	13.8	6798	22	AAI81954	Streptomyces sp. C
41	37.6	13.8	6798	22	AAI81954	Streptomyces sp. C
c	37.6	13.8	6798	22	AAI81954	Nucleotide sequenc
42	37.2	13.6	665	22	AAI81954	Partial nucleotide
43	37	13.6	30001	18	AAI81954	Total DNA sequence
c	36.8	13.5	58857	21	AAI81954	S. aureofaciens DN
45	36.8	13.5	58857	21	AAI81954	Nucleotide sequenc

ALIGNMENTS

RESULT 1	
AAZ34955	
ID	AAZ34955 standard; cDNA; 273 BP.
XX	
AC	AAZ34955;
XX	
DT	28-FEB-2000 (first entry)
XX	
DE	Corn acid triacylglycerol lipase cDNA.
XX	
KW	Triacylglycerol lipase; corn; maize; fatty acid; seed oil;
KW	vegetable oil; transgenic plant; ss.
XX	
OS	Zea mays.
XX	
PH	Key
FT	CDS
FT	2..272
XX	/*tag= a
PN	WO9955883-A2.
XX	
PD	04-NOV-1999.
XX	
PF	29-APR-1999; 99WO-US09280.
XX	
PR	30-APR-1998; 98US-0083688.
XX	
PA	(DUPO) DU FONT DE NEMOURS & CO E I.
XX	
PI	Cahoon EB, Cahoon RE, Kinney AJ, Rafalski JA;
XX	
DR	WPI; 2000-062036/05.

DR P-PSDB; AAY32306.
XX Novel plant triacylglycerol lipase polynucleotides used to alter the
PT level of the enzyme in transgenic plants
XX
XX Claim 2; Page 43; 65pp; English.
XX
XX This is the nucleotide sequence of the cDNA insert in clone
CC p0102.cnlcm37r encoding a portion (see AAY32306) of corn acid
CC triacylglycerol lipase (TAGL). The clone was isolated from a corn
CC leaf cDNA library. Novel acid and neutral TAGL polypeptides
CC (see AAY32301-17) and polynucleotides (see AAY34950-66) from corn,
CC catalpa, rice, soybean and wheat tissues are disclosed. The enzymes
CC may be prepared recombinantly and used to raise antibodies, which
CC are used for detecting the enzymes in situ in cells or in vitro in
CC cell extracts. The polynucleotides may be used to create transgenic
CC plants in which the TAGL levels are present at higher or lower levels
CC than normal, or in cell types or developmental processes where they are
CC not normally found. This would alter the level of triacylglycerol and
CC cholesterol esters found in those cells. Accumulation of fatty acids
CC with unusual structures may be a positive phenotype in plants used for
CC foods. In addition, it may be desirable to eliminate expression of TAGL
CC genes for certain applications. TAGL enzymes may also be useful for the
CC processing of plant seed oils and for the development of novel seed
CC oils. The TAGL enzymes can also be used as targets to facilitate the
CC design and/or identification of inhibitors of those enzymes that may be
CC useful as herbicides. This is desirable because inhibition of the
CC activity of either of the enzymes could lead to an inhibition of plant
CC growth. The polynucleotides also serve as a source of probes and
CC primers, which are useful for genetic mapping, as markers for traits
CC linked to those genes, and to isolate homologous sequences from other
CC species.
XX
SQ Sequence 273 BP; 46 A; 87 C; 79 G; 54 T; 7 other;

Query Match 97.4%; Score 266; DB 21; Length 273;
Best Local Similarity 100.0%; Pred. No. 7.8e-34;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ctctctntgacgcttgcgttctgactctgactggaactggtcctgggtag 60
Db 1 ctctctntgacgcttgcgttctgactctgactggaactggtcctgggtag 60
Qy 61 tcaacgacctgcggccatggtgcgactctgctgcaaacagacgagcctcact 120
Db 61 tcaacgacctgcggccatggtgcgactctgctgcaaacagacgagcctcact 120
Qy 121 acgtcgacactccatgggagcgtggtggtggtggtggtggtggtggtggtg 180
Db 121 acgtcgacactccatgggagcgtggtggtggtggtggtggtggtggtggtg 180
Qy 181 tgagcagctgaaatcccgagcgtggtggtggtggtggtggtggtggtggtg 240
Db 181 tgagcagctgaaatcccgagcgtggtggtggtggtggtggtggtggtggtg 240
Qy 241 nccccaatnggaatcctggtggtggtggtggtggtggtggtggtggtggt 273
Db 241 nccccaatnggaatcctggtggtggtggtggtggtggtggtggtggtggt 273

RESULT 2
AAZ34958
ID AAZ34958 standard; cDNA; 1718 BP.
XX
XX
AC AAZ34958;
XX
XX 28-FEB-2000 (first entry)
XX
XX Soybean acid triacylglycerol lipase cDNA.
XX
XX Triacylglycerol lipase; soybean; fatty acid; seed oil;
KW vegetable oil; transgenic plant; ss.
KW

XX Glycine max.
OS
XX
FH Key Location/Qualifiers
FT CDS 184..1416
FT /*tag= a
XX
XX W09955883-A2.
XX
XX 04-NOV-1999.
XX
XX 29-APR-1999; 99WO-US09280.
XX
XX 30-APR-1998; 98US-0083688.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Caboon EB, Cahoon RE, Kinney AJ, Rafalski JA;
XX
XX WPI; 2000-062036/05.
XX
XX P-PSDB; AAY32309.
XX
XX Novel plant triacylglycerol lipase polynucleotides used to alter the
XX level of the enzyme in transgenic plants
XX
XX Claim 2; Page 48; 65pp; English.
XX
XX This is the nucleotide sequence of a contig assembled from the
XX entire cDNA insert in clone ssl.pk0022.al and a portion of the
XX cDNA insert in clone sop3c.pk004.n3 encoding the entire soybean
XX acid triacylglycerol lipase (TAGL) (see AAY32309). The clones were
XX isolated from soybean developing pod and seedling (5-10 day post
XX germination) cDNA libraries. Novel acid and neutral TAGL polypeptides
XX (see AAY32301-17) and polynucleotides (see AAY34950-66) from corn,
XX catalpa, rice, soybean and wheat tissues are disclosed. The enzymes
XX may be prepared recombinantly and used to raise antibodies, which
XX are used for detecting the enzymes in situ in cells or in vitro in
XX cell extracts. The polynucleotides may be used to create transgenic
XX plants in which the TAGL levels are present at higher or lower levels
XX than normal, or in cell types or developmental processes where they are
XX not normally found. This would alter the level of triacylglycerol and
XX cholesterol esters found in those cells. Accumulation of fatty acids
XX with unusual structures may be a positive phenotype in plants used for
XX foods. In addition, it may be desirable to eliminate expression of TAGL
XX genes for certain applications. TAGL enzymes may also be useful for the
XX processing of plant seed oils and for the development of novel seed
XX oils. The TAGL enzymes can also be used as targets to facilitate the
XX design and/or identification of inhibitors of those enzymes that may be
XX useful as herbicides. This is desirable because inhibition of the
XX activity of either of the enzymes could lead to an inhibition of plant
XX growth. The polynucleotides also serve as a source of probes and
XX primers, which are useful for genetic mapping, as markers for traits
XX linked to those genes, and to isolate homologous sequences from other
XX species.
XX
SQ Sequence 1718 BP; 492 A; 353 C; 344 G; 529 T; 0 other;

Query Match 36.0%; Score 98.2; DB 21; Length 1718;
Best Local Similarity 63.0%; Pred. No. 3e-14;
Matches 148; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
Qy 24 cagctctactggaactggtcctctgggtagctgacctggttagtcaacgacctgccgacctggtc 83
Db 619 caggcctattggaattggtcttggatgaactgtctctctatgattccctcggtgttt 678
Qy 84 gacttcgtctgcaaacagacgagccgaggaagcctcactacgtcggaactccatccatgggagc 143
Db 679 aattatgtgttcagccaaacgagggcagaagatcaattacgttggccattcattgggaact 738
Qy 144 ctggtggtcgctggcggtctcttcgagggcggtggtggtggtggtggtggtggtggtggtg 203
Db 739 ttggtgagctttggcatcctctcgggaaggaataattggttaccctccgactgaaacagcagcc 798

QY 204 ctgctcacccggtgacctacgtcgccacatnaaccccccaatngaatctg 258
 Db 799 ttgtgagccctatagccctatttaagccacatgaatacacagcacttggtgtg 853

RESULT 3
 ID ABL24956/c
 XX ABL24956 standard; DNA; 2608 BP.
 AC ABL24956;
 XX 26-MAR-2002 (first entry)
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 26341.
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 XX Drosophila melanogaster.
 OS Drosophila melanogaster.
 PN WO200171042-A2.
 XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US09231.
 PF 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EM;
 PI WPI; 2001-656860/75.
 DR New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX Claim 1; SEQ ID NO 26341; 21pp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 2608 BP; 755 A; 542 C; 558 G; 753 T; 0 other;

Query Match 20.9%; Score 57; DB 23; Length 2608;
 Best Local Similarity 58.8%; Pred. No. 0.00016;
 Matches 134; Conservative 0; Mismatches 88; Indels 6; Gaps 2;

QY 24 cagctcactggaactggtctggtgacctgacctgtagtcaacgacctgcccggcctggtc 83
 Db 2546 CAGATCTCTGGAACCTCAGCTGGAAGAGATCGGCATGTACGACGTTCGGCGATGATT 2487
 QY 84 gacttcgtctcaaacagaccgcccagaagcct---cactacgtcgacactccatgggg 140
 Db 2486 GATTATGCTCTGGCAAGACTGGACAGCAGCAGGTGCAATATGTGGGCCACTCCAGGGT 2427
 QY 141 acgcgtggtggcgtggcgccctcttcgagggccgggtggtga---gccagctgaaatcc 197
 Db 2426 ACCACTGCTATCTGTTGATGGTGTGTCGGAGAGACCCGAGTACATGACAAAGATCAAGTCC 2367

QY 198 ggcgcgtgctcacccggtggcctacctcgccacatnaaccccc 245
 Db 2366 GCCATCTTCTGGGACCAGCCGCTATATGGCAACATGAAGAGCCCC 2319

RESULT 4
 ID AAZ34956
 XX AAZ34956 standard; cDNA; 1483 BP.
 AC AAZ34956;
 XX 28-FEB-2000 (first entry)
 DE Rice acid triacylglycerol lipase cDNA.
 XX Triacylglycerol lipase; rice; fatty acid; seed oil;
 KW vegetable oil; transgenic plant; ss.
 XX Oryza sativa.
 OS Oryza sativa.
 FH Key Location/Qualifiers
 FT CDS 32..1264
 FT /*tag= a
 XX WO955883-A2.
 XX 04-NOV-1999.
 XX 29-APR-1999; 99WO-US09280.
 XX 30-APR-1998; 98US-0083688.
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 PI Cahoon EB, Cahoon RE, Kinney AJ, Rafalski JA;
 XX WPI; 2000-062036/05.
 DR P-PSDB; AAY32307.
 XX Novel plant triacylglycerol lipase polynucleotides used to alter the
 PT level of the enzyme in transgenic plants -
 XX Claim 2; Page 44; 65pp; English.

XX This is the nucleotide sequence of the cDNA insert in clone
 CC rlr72.ph0015.b2 encoding the entire rice acid triacylglycerol
 CC lipase (TAGL) (see AAY32307). The clone was isolated from a rice
 CC leaf cDNA library. Novel acid and neutral TAGL polypeptides
 CC (see AAY32301-17) and polynucleotides (see AAZ34950-66) from corn,
 CC Catalpa, rice, soybean and wheat tissues are disclosed. The enzymes
 CC may be prepared recombinantly and used to raise antibodies, which
 CC are used for detecting the enzymes in situ in cells or in vitro in
 CC cell extracts. The polynucleotides may be used to create transgenic
 CC plants in which the TAGL levels are present at higher or lower levels
 CC than normal, or in cell types or developmental processes where they are
 CC not normally found. This would alter the level of triacylglycerol and
 CC cholesterol esters found in those cells. Accumulation of fatty acids
 CC with unusual structures may be a positive phenotype in plants used for
 CC foods. In addition, it may be desirable to eliminate expression of TAGL
 CC genes for certain applications. TAGL enzymes may also be useful for the
 CC processing of plant seed oils and for the development of novel seed
 CC oils. The TAGL enzymes can also be used as targets to facilitate the
 CC design and/or identification of inhibitors of those enzymes that may be
 CC useful as herbicides. This is desirable because inhibition of the
 CC activity of either of the enzymes could lead to an inhibition of plant
 CC growth. The polynucleotides also serve as a source of probes and
 CC primers, which are useful for genetic mapping, as markers for traits
 CC linked to those genes, and to isolate homologous sequences from other
 CC species.

XX Sequence 1483 BP; 375 A; 334 C; 344 G; 430 T; 0 other;

XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX Claim 1; SEQ ID NO 24313; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 5137 BP; 1422 A; 1190 C; 1101 G; 1424 T; 0 other;

Query Match 17.9%; Score 48.8; DB 23; Length 5137;
Best Local Similarity 56.3%; Pred. No. 0.014;
Matches 130; Conservative 0; Mismatches 95; Indels 6; Gaps 2;
Qy 21 ttacagctctactggaactgtctgtggtgacgctgtagtcaacgacctccgagccatg 80
Db 1743 tatcaggagttctggacttcacctccacgagatgggcaagtacgacctccgagccat 1802
Qy 81 gtcagctctgcgtcaaacagacagg---ccagaagcctcactacgtcgacactccatg 137
Db 1803 atcgactacatccgagacaggggatacagcaggtccactacatcgctcactccacg 1862
Qy 138 gggacgctgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 197
Db 1863 ggaacagcattcttctgt 1922
Qy 198 gggcgcgtg---ctcagcgggtggtggtggtggtggtggtggtggtggtggtggtggtggt 245
Db 1923 tccatgcacgcccgtggcgccatcgctacacacacacacacacacacacacacacacac 1973

RESULT 13
AAZ87302
ID AAZ87302 standard; DNA; 2787 BP.
XX AAZ87302;
XX 05-JUN-2000 (first entry)
XX S. venezuelae macrolide biosynthetic gene pikD, SEQ ID NO:40.
XX Desosamine biosynthesis; macrolide; polyketide; methymycin; pikromycin;
XX neomethymycin; narbomycin; polyhydroxyalkanoate monomer synthase;
XX biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma,
XX chronic obstructive pulmonary disease; respiratory inflammation;
XX hypercholesterolaemia; crop protection agent; ds.
XX Streptomyces venezuelae ATCC15439.
XX Key Location/Qualifiers

CDS 1..2787
FT FT /*tag= a
XX /product= "PikD"
PN WO200000620-A2.
XX 06-JAN-2000.
XX 25-JUN-1999; 99WO-US14398.
XX 26-JUN-1998; 98US-0105537.
XX (MINU) UNIV MINNESOTA.
PI Sherman DH, Liu H, Xue Y, Zhao L;
XX WPI; 2000-160679/14.
DR P-PSDB; AAY77197.
XX Desosamine and macrolide biosynthetic gene clusters, useful for, e.g.
XX synthesis of methymycin and pikromycin -
XX Claim 15; Page 431-432; 438pp; English.
XX The invention relates to an isolated and purified nucleic acid segment
XX comprising a desosamine biosynthetic gene cluster, a fragment or its
XX biologically active variant, where the nucleic acid sequence is not
XX derived from the eryC gene cluster of Saccharopolyspora erythraea or
XX Streptomyces antibioticus. The invention also relates to a macrolide
XX biosynthetic gene cluster, or fragments thereof. The macrolide
XX biosynthetic gene cluster encodes proteins which synthesize methymycin,
XX pikromycin, neomethymycin, narbomycin or a combination of these
XX compounds. Recombinant or augmented cells comprising the desosamine
XX and/or macrolide biosynthetic gene clusters are useful for the
XX production of biologically active macrolides. The macrolide biosynthetic
XX proteins are useful for synthesis of methymycin, pikromycin,
XX neomethymycin and narbomycin. The alternative termination of polyketide
XX synthesis may be useful to prepare novel antibiotics and
XX polyhydroxyalkanoate (PHA) monomers. The compounds produced by the
XX recombinant host cells are useful as biopolymers, e.g., in packaging or
XX biomedical applications, to engineer PHA monomer synthases or to prepare
XX biologically active agents, such as chemotherapeutics,
XX immunosuppressants, agents to treat asthma, chronic obstructive pulmonary
XX disease as well as other diseases involving respiratory inflammation,
XX cholesterol-lowering agents or macrolide-based antibiotics which are
XX active against a variety of organisms, e.g., bacteria, including
XX multi-drug resistant pneumococci and other respiratory pathogens, as well
XX as viral parasitic pathogens, or as crop protection agents (e.g.,
XX fungicides or insecticides) via expression of polyketides in plants.
XX Sequences AAZ87295-287302 represent macrolide biosynthetic genes from
XX Streptomyces venezuelae ATCC 15439, which encode proteins
XX AAY77190-Y77197.
SQ Sequence 2787 BP; 370 A; 1104 C; 972 G; 341 T; 0 other;

Query Match 14.9%; Score 40.6; DB 21; Length 2787;
Best Local Similarity 51.1%; Pred. No. 1.1;
Matches 94; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
Qy 49 atgacctgtgtgtagtcaacgacctgcccgcacatggtgacctgtgctcaaacagacgccc 108
Db 1592 aggccagggccggcgtctccagcggggcccggaacacctgctgcgagcggcgaac 1651
Qy 109 agagcctcactagctcgacactccatggagacgtggtggtggtggtggtggtggtggtggt 168
Db 1652 agatcctcagggcgtgctgctggtggtggtggtggtggtggtggtggtggtggtggtggt 1711
Qy 169 agggcggggtggtgagccagctgaaatccgcggcgtcgtcagccggtggtggtggtggtggtggt 228
Db 1712 tggctcgtctccacgcgacggcgtcgcggcggtggtggtggtggtggtggtggtggtggtggt 1771
Qy 229 ncca 232

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 02:13:07 ; Search time 1843.33 Seconds
(without alignments)
3099.252 Million cell updates/sec

Title: US-09-699-652-11
Perfect score: 273
Sequence: 1 ctctctcgtgcacgttcgn.....tcttggtggccaangcggt 273

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	GenEmbl.*
1:	gb_ba.*
2:	gb_htg.*
3:	gb_in.*
4:	gb_om.*
5:	gb_ov.*
6:	gb_pat.*
7:	gb_ph.*
8:	gb_pl.*
9:	gb_pr.*
10:	gb_ro.*
11:	gb_sts.*
12:	gb_sy.*
13:	gb_un.*
14:	gb_vi.*
15:	em_ba.*
16:	em_fun.*
17:	em_hum.*
18:	em_in.*
19:	em_mu.*
20:	em_om.*
21:	em_or.*
22:	em_ov.*
23:	em_pat.*
24:	em_ph.*
25:	em_pl.*
26:	em_ro.*
27:	em_sts.*
28:	em_un.*
29:	em_vi.*
30:	em_htg_hum.*
31:	em_htg_inv.*
32:	em_htg_other.*
33:	em_htgo_inv.*

ALIGNMENTS

RESULT	1	AP003920/c
LOCUS	AP003920	140319 bp DNA linear HTG 12-JUL-2001
DEFINITION	Oryza sativa chromosome 8 clone OJ1789_C07, *** SEQUENCING IN PROGRESS ***, in ordered pieces.	
ACCESSION	AP003920	
VERSION	AP003920.1	GI:14701606
KEYWORDS	HTG; HTGS_PHASE2.	
SOURCE	Oryza sativa (cultivar:Nipponbare) DNA, clone:OJ1789_C07.	
ORGANISM	Oryza sativa	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.	
AUTHORS	1 (bases 1 to 140319)	
TITLE	Sasaki,T., Matsumoto,T. and Yamamoto,K.	
JOURNAL	Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, BAC clone:OJ1789_C07	
REFERENCE	Published Only in Database (2001) In press	
AUTHORS	2 (bases 1 to 140319)	
TITLE	Sasaki,T., Matsumoto,T. and Yamamoto,K.	
JOURNAL	Direct Submission	
REFERENCE	Submitted (11-JUL-2001) Takuji Sasaki, National Institute of Agrobiological Resources, Rice Genome Research Program; Kannondai	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
-----	-----	-----	-----	-----	-----	-----

This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is MAC12 and the 3' clone is F18022.

FEATURES

source

Location/Qualifiers
1..65465

/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="5"
/clone="MUA22"
/clone_lib="Mitsui P1"

complement(1..95)

/notes="CDS is reported in Acc# AB005230"

gene_id:MAC12.1

unknown protein"

/number=8

/evidence=not_experimental

complement(191..298)

/notes="CDS is reported in Acc# AB005230"

gene_id:MAC12.1

unknown protein"

/number=7

/evidence=not_experimental

complement(422..520)

/notes="CDS is reported in Acc# AB005230"

gene_id:MAC12.1

unknown protein"

/number=6

/evidence=not_experimental

complement(614..680)

/notes="CDS is reported in Acc# AB005230"

gene_id:MAC12.1

unknown protein"

/number=5

/evidence=not_experimental

complement(756..852)

/notes="CDS is reported in Acc# AB005230"

gene_id:MAC12.1

unknown protein"

/number=4

/evidence=not_experimental

complement(934..1084)

/notes="CDS is reported in Acc# AB005230"

gene_id:MAC12.1

unknown protein"

/number=3

/evidence=not_experimental

complement(1184..1398)

/notes="CDS is reported in Acc# AB005230"

gene_id:MAC12.1

unknown protein"

/number=2

/evidence=not_experimental

complement(1474..1678)

/notes="CDS is reported in Acc# AB005230"

gene_id:MAC12.1

unknown protein"

/number=1

/evidence=not_experimental

join(2737..2918,3227..3325,3410..3518,3687..3758,3993..4112)

/notes="gene_id:MUA22.2

similar to unknown protein

sp|P23438"

/codon_start=1

/evidence=not_experimental

/protein_id="BAB08282.1"

/db_xref="GI:9757784"

/translation="MAVAVAKLLISAMAVFMLYSASFATSEVPFPMVYVHKATLNRLKS
GAEKVSVDIYNQSSSADVTLDNSDKKTFEVVNGNTSKSWEKLDAGGLSHSI
ELEAKVKGYPGAVVVFRIPTKPAQEQAYSTPLPLDLADKPPTKPLDVRLLAKY
GSLVSVISMWVCFIYLATPKSNVSKASSKKR"
6333..6408

tRNA

CDS

/note="codon recognized: CUU; gene_id:MUA22.3"
/product="tRNA-Lys"
/evidence=not_experimental
complement(join(6918..6978,7067..7191,7318..7532,
7659..7702,7780..8141,8504..8824))

/note="gene_id:MUA22.4"

/codon_start=1

/evidence=not_experimental

/product="mitochondrial phosphate translocator"

/protein_id="BAB08283.1"

/db_xref="GI:9757785"

/translation="MESPKNSLIPSLYSSSSSPRFLDLQVLNSNSNAFEKSPSPA
PRSSPTSMISRNKFLIASPTGKGIEMSPAFYAACFTGGILSCGLTHMTVPLDLV
KCNMQIDPAKYKISISGFGILLKEQGVKGFRRGVTLLGYSQAQCKFGFYFYFKKT
YSLDAGPEYAKYKTLIYLAGSASAEIADIALCPFEAVKRVQTOPGFARGMSDGFPP
KFTKSEGYGLYKGLAPLMGROIPTYMMKFASPETIVEMLYKVAIINPKSECSKGLQL
GVSFAGYVAGVFCIAIVSHPADNLVSLFNKAGTAVGDAAVKIGMGLFTRGLPLRIV
MIGTLTGAQWGLYDAKFVFGVLTGGVAPAPAIAATEAKA"
9398..11038

/note="gene_id:MUA22.5"

ref|NP_057085.1

similar to unknown protein"

/codon_start=1

/evidence=not_experimental

/protein_id="BAB08284.1"

/db_xref="GI:9757786"

/translation="MSLSQNAKSKGIKREELKQYEDVEDEEIGSDDDLTRGKREK
TEKQKLESELVEMKLENLIFGSLYPTFGKEEEDGSAHFVDRSAVRQIPDY
EDGDDDEELSDDEENGQVVAIRKEGAWEDEEKQINVDIASVNRKLRKEENGLI
SGSEYIARLRAHAKLNPQTDWAPDSQIVDGESSDDDDTDQGGVDDILRTNEDLVVK
RNTGNKLCAGLEYSKLVANADPSNPINSVHFHQAQLLTAGLDRRLRFQIDCK
RNTKIOSIFLEDCPIRKAFLPNGSOVIVSGRRKFFYSFOLKAKFDKIGPLVGREK
SLEYFEVSQDSNTIAFVNEGYYLLYSTKTELIGTLKMGNSVRSIAFSEDGKHLSS
GQGQVYVLDLRTMKLYKGVDEGTCGTSLSLNGALFASGTDGRLGIVNIYKKSERV
GKKRKPIKTVNLTSLKIDFMKNDAQILAIVSTMNKNKSVKLVHVPSLTVFSNMPPPN
STMHPYPCLDSPGSGFMANGAGKVLVLYKLLHYQNA"

join(12002..12391,12485..12562,12671..12711,12820..12974,
13144..13293,13467..13562,13634..13750,13832..13951,
14029..14121,14202..14324,14407..14523,14606..14668,
14766..14852)

/note="gb|AA063104.1

gene_id:MUA22.6"

/codon_start=1

/evidence=not_experimental

/product="lysine-sensitive aspartate kinase"

/protein_id="BAB08285.1"

/db_xref="GI:9757787"

/translation="MASLIQLYGVKTPGLALSSKRLEFASKGACFSVTLPSSSAVFRDV
EHCNRNIGLRVSCALRVLDLQKREPETCDSSGTGKELTCVMKFGSSVSASERMKEV
ANLILSPDPRPVIVLSAMGKTTNKLKAGEKAVTCGTVNVESTIELSFTKELHLRTA
HELGVETVTEKHEGLHOLLKGISMMKELTLETRDYLVSFGECSMTRFLPSAYLNKIG
HKARQDAFEIGITDDFTNADILENTYPAVSKTLVGDMSKENVAVPVVGYLGKGNR
SCAITTLGRGSDLTATPICKALGLEIRIQMKVDVGLTCDPNIIYPGAQSVPLTDFE
AAELAYFGAQLHLPLSMRPARDGIPVRVKNNSNPATGTVITRSRDSKAVLTSLVL
KRWVMDLIASTRMLQGYGLAKVFTFEDLIGISVDVAVTSEVLSITLDPKALMGRE
LIQVNELDNVELEKIAVVKLLQRRSIIISLIGNQKSSLLILEKVFQVFRSNGVNVQ
MISQASKVNIISLVNDEEAECQVRALHSAFFETDP"
17774..18196

/note="gene_id:MUA22.7"

/codon_start=1

/evidence=not_experimental

/product="glutaredoxin-like protein"

/protein_id="BAB08286.1"

/db_xref="GI:9757788"

/translation="MQYKTRGSLSYNNNSKYNMNNVFPSETLAKIESMAENAVV
IFSVSTCCMCHAIKRLFRGMGSPAVHELDLLPYGVIEIRALLRLLCSSGGATSPGA
LPVVFIGKMGVAMERVMASHINGSLVPLLKDGALWL"
complement(19124..21115)

/note="gb|AA019552.1

gene_id:MUA22.8

similar to unknown protein"

/codon_start=1

/evidence=not_experimental

1_d3 (L77000), 3_c5 (AC000599), 3_h2, 3_h6 (AC000600), 2_e3 (AC000598).	
FEATURES	
source	Location/Qualifiers
	1. 82393
	/organism="Drosophila melanogaster"
	/db_xref="taxon:7227"
	/chromosome="3R"
	/map="87F1-87F2"
	/clone="P1 DS04219 (D36)"
BASE COUNT	23743 a 17464 c 17263 g 23923 t
ORIGIN	
	Query Match 20.9%; Score 57; DB 3; Length 82393;
	Best Local Similarity 58.8%; Pred. NO. 0.18;
	Matches 134; Conservative 0; Mismatches 88; Indels 6; Gaps 2;
QY	24 cagctctactggaactggtcctctgggatgacctggtagtcaacgacctcccgcccatggtc 83
Db	76093 CAGATCTTCTGGAACCTTCACGTGGACGAGATCGGCATGTACGACGTTCCGGCGATGATT 76152
QY	84 gacttcgtcgtcaaacagaccgcccagaaagcct---cactacgtcggacatcccatggg 140
Db	76153 GATTATGCTCTTGCAAAAGACTGGACAGCAGCAGAGGTGCAATATGTGGGCCATCCAGGGT 76212
QY	141 acgctggtggtgcgtggcgccctctcggaggccgggtggtga---gccagctgaaatcc 197
Db	76213 ACCACTGTCTATCTGTGTGATGCTGCGGAGAGACCCGAGTACATGACAAGATCAAGTCC 76272
QY	198 gggcgctactcacgcccgggtggcctacctcgcncacatnaaccccc 245
Db	76273 GCCATCTTCTGGACACCGCCCTATATGGCAACATGAAGAGCCCC 76320
RESULT	7
AC008360/c	
LOCUS	AC008360 194672 bp DNA linear INV 30-JUN-2001
DEFINITION	Drosophila melanogaster, chromosome 3R, region 87E-87F, BAC clone BACR06H18, complete sequence.
ACCESSION	AC008360
VERSION	AC008360.3 GI:14578131
KEYWORDS	HTG.
SOURCE	fruit fly.
ORGANISM	Drosophila melanogaster
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
AUTHORS	1 (bases 1 to 194672) Celnikier,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H., Holt,R.A., Evans,C.A., Gocayne,J.D., Ananidis,P.G., Brandon,R.C., Rogers,Y., An,H., Baldwin,D., Bazon,J., Beeson,K.Y., Busam,D.A., Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D., Ferriera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A., Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J., Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattel,B., Moshrefi,A., McIntosh,T.C., Moy,M., Murphy,S., Nelson,C., Nelson,K.A., Nunoo,J., Pacleb,J., Paragas,V., Park,S., Patel,S., Patel,S., Pfeiffer,B., Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F., Stapleton,M., Strong,R., Svirska,R., Tector,C., Williams,S.M., Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
TITLE	Sequencing of Drosophila chromosome 3R, region 87E-87F
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 194672) Celnikier,S.E., Aghayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhoff,C., Chabme,N., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R.A., George,R.A., Harris,N.L., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E., Svirska,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.

TITLE Direct Submission
JOURNAL Submitted (02-AUG-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT On Jun 30, 2001 this sequence version replaced gi:5706740.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121 Berkeley, CA 94720
This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email to bdg@fruitfly.berkeley.edu.
FEATURES
source
1. 194672
/organism="Drosophila melanogaster"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="3R"
/map="87E-87F"
/clone="BACR06H18 (D979)"
/clone.lib="RPI-98 (Roswell Park Cancer Institute Drosophila melanogaster BAC library, partial EcoRI in pBACe3.6)"
BASE COUNT 56534 a 41027 c 40879 g 56232 t
ORIGIN
Query Match 20.9%; Score 57; DB 3; Length 194672;
Best Local Similarity 58.8%; Pred. No. 0.16;
Matches 134; Conservative 0; Mismatches 88; Indels 6; Gaps 2;
QY 24 cagctctactgaactggtctctggatgacgtcgtgtgtagtcaacgacctccgcccgtatgctc 83
Db 72433 CAGATCTTCTGGAACCTCAGCTGGACAGATGCGCATGTACGACGTTCCGGCGATGATT 72374
QY 84 gacttcgtcgtcaaacacagccggcgaagaagcct---cactacgtcggacatccatccatggg 140
Db 72373 GATTATGCTCTGGCAAGACTGGACAGCAGCAGGTCGTAATATGTGGCCACTCCAGGGT 72314
QY 141 acgctcgttggtcgtcggcgtcttcgagagggcgggtggtga---gccagctgaatccc 197
Db 72313 ACCACTGCTATCTGCTGATGCTGTCGAGAGACCCGAGTAGTCAATGACAAATCAAGTTC 72254
QY 198 gcggcgctcgtcacgcgggtggtcactcctcgcacatnaacncccc 245
Db 72253 GCCCATCTCTGGGACGACGCCCTATATGCGCAACATGAAGAGCCCC 72206
RESULT 8
LOCUS AE003699/c
DEFINITION Drosophila melanogaster genomic scaffold 142000013386035 section 24 of 105, complete sequence.
ACCESSION AE003699 AE002708
VERSION AE003699.2 GI:10726505
KEYWORDS HTG.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 225668)
AUTHORS Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor Miklos G.L., Abril J.F., Aghayani A., An H.J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y.,

Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Bottier P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W., Fowler C., Gabriellian A.E., Garg N.S., Gelbart W., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.H., Idegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Neison K.A., Nixon K., Nusskern D.R., Pacieb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M. and Venter J.C.
The genome sequence of Drosophila melanogaster
Science 287 (5461), 2185-2195 (2000)
2 (bases 1 to 225668)
ADAMS M.D., CELNIKER S.E., GIBBS R.A., RUBIN G.M. and VENTER, C.J.
Direct Submission
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA
On Oct 9, 2000 this sequence version replaced gi:7293726.
FEATURES
source
1. 225668
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="3R"
join(<14264. .14296,14382. .14587,14883. .14971,15275. .15478,15629. .>17028)
/gene="timeout"
/product="CT24346"
/db_xref="FLYBASE:FBgn0008148"
/db_xref="FLYBASE:FBgn0038118"
/evidence=not_experimental
<14264. .>17028
/gene="timeout"
/note="CG8148"
/map="07D13-87E1"
/db_xref="FLYBASE:FBgn0008148"
/db_xref="FLYBASE:FBgn0038118"
/evidence=not_experimental
join(14264. .14296,14382. .14587,14883. .14971,15275. .15478,15629. .17028)
/gene="timeout"
/note="timeout gene product"
/codon_start=1
/db_xref="FLYBASE:FBgn0008148"
/db_xref="FLYBASE:FBgn0038118"
/evidence=not_experimental
/protein_id="AAF54908.1"
/db_xref="GI:7293727"
/translation="MVEKRGEVAAVAKLFRIFQQVSVVERDVHVEELRLRLAIFVVRKF
RVNPTDKPIYAEILLDFYKGRREANELESYGDYEAETKGAWTEQESLRFEEENQ
ENPTPTDKVIDWLDNDVKTNRVTMLKELGLLFKAPTKRSTKSAQSGKNWQ
EEDDELRLSDVQRIEPCDLERLVNEFAERSKQIITKRLQIHLIAIDKSLIPAKKG
RGDKPKKDVMEGEGEFDFPEFMDGQYKKPKKVKQRMVPTLDVGTIR
ALIGQVSEKYSQSALEWLQCLQDASEDEVEDDGVPLPLPLMENQKNAMEDGDFQ

REFERENCE 2 (bases 1 to 105659)
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-2001) Takuji Sasaki, National Institute of
Agrobiological Resources, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
COMMENT The nucleotide sequence of this BAC clone was generated by
combining Monsanto and RGP-Japan sequencing data.
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
Location/Qualifiers
source
1. 105659
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="6"
/clone="QJ1628.G11"
BASE COUNT 28723 a 24017 c 23319 g 29500 t 100 others
ORIGIN

Query Match 19.7%; Score 53.8; DB 2; Length 105659;
Best Local Similarity 65.3%; Pred. No. 0.66;
Matches 79; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 21 ttctagctctactgaactgtctctggatgacctggttagtcaacgacctccggccatg 80
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61824 TTTCAGGCTTACTGGGAATGTCGTGGGACCAACTTCTTCTAGCATCTTCTCGAGTG 61765
QY 81 gtcgactctgtctcaacagaccgcccagagcctcactcgtcgacacctccatggg 140
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61764 CTGCACTTGGCTATGATCACACAGGAGAAATCCACTATATCGGTCTCCCTGGTG 61705
QY 141 a 141
Db 61704 A 61704

RESULT 10
AC020487
LOCUS Drosophila melanogaster, 50402 bp DNA linear HTG 03-JAN-2000
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
pieces.
ACCESSION AC020487
VERSION AC020487.1 GI:6664410
KEYWORDS HTG; HTGS_PHASE2.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 50402)
AUTHORS Adams, M. and Venter, J.C.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT This sequence was identified as CDM:10213375 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
Location/Qualifiers

source 1. 50402
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
BASE COUNT 13641 a 11696 c 11744 g 13321 t
ORIGIN

Query Match 18.8%; Score 51.4; DB 2; Length 50402;
Best Local Similarity 57.6%; Pred. No. 2;
Matches 129; Conservative 0; Mismatches 89; Indels 6; Gaps 2;
QY 27 ctctactggaactgtctctggatgacctggttagtcaacgacctccggccatggtcgac 86
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 28494 GCCTTCTGCGATTTCAGCTGGCAGCATGGCATCTACGACCTGCCGCCATCTGGAT 28553
QY 87 ttctg---tcgtcaaacagaccgcccagagcctcactcgtcgacacctccatggggagcg 143
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 28554 TACGTGCTGGGAGGACGAATCGCCGCGACTGTCAGTCGGGCACCTCGCAGGCACC 28613
QY 144 ctggtggcgctggcgccctcttcggagggcggtggtagcgacgtgaaatccc---gcg 200
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 28614 ACCGTTCTCTGCTGCTCTCGCAGCGCGGAGTACATGCGGGTTCGCCAATGCC 28673
QY 201 gogctgtcaacgcccgtggttcctacctcgnccatnaaacnccc 244
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 28674 GCCCTGTTGGTCCGGTGCGATTCTCCAGCACCTGAGCAGTCC 28717

RESULT 11
AC009460/c
LOCUS Drosophila melanogaster, 159455 bp DNA linear INV 14-MAR-2001
DEFINITION Drosophila melanogaster, chromosome 2R, region 60B-60C, BAC clone
BACR04P18, complete sequence.
ACCESSION AC009460
VERSION AC009460.4 GI:13324757
KEYWORDS HTG.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 159455)
AUTHORS Celisner, S.E., Adams, M.D., Kronmiller, B., Tyler, D., Wan, K.H.,
Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C.,
Rogers, Y., An, H., Baldwin, D., Banzon, J., Beeson, K.Y., Busan, D.A.,
Carlson, J.W., Center, A., Champs, M., Davenport, L.B., Dietz, S.M.,
Dodson, K., Dorsett, V., Doup, L.E., Doyle, C., Dresnek, D., Farfan, D.,
Ferriera, S., Frise, E., Galle, R.F., Garg, N.S., George, R.A.,
Gonzalez, M., Houch, J., Hoskins, R.A., Hostin, D., Howland, T.J.,
Ibegwam, C., Jalali, M., Kruse, D., Li, P., Mattei, B., Moshrefi, A.,
McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunoo, J.,
Pacleb, J., Paragas, V., Park, S., Patel, S., Pfeiffer, B., Scheeler, F.,
Phouanavong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F.,
Stapleton, M., Strong, R., Svirskas, R., Tector, C., Williams, S.M.,
Zaveri, J.S., Smith, H.O., Rubin, G.M. and Venter, J.C.
Sequencing of Drosophila chromosome 2R, region 60B-60C
Unpublished
2 (bases 1 to 159455)
AUTHORS Celisner, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,
Buenhoff, C., Champs, M., Chavez, C., Chew, M., Ciesiolka, L.,
Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,
Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K., Kearney, L.,
Kim, E., Lee, B., Lewis, S., Li, P., Lomotan, M.A., Mazda, P.,
Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S.,
Pfeiffer, B., Poon, L., Sequeira, A., Sethi, H., Snir, E.,
Svirskas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and
Rubin, G.M.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT On Mar 14, 2001 this sequence version replaced gi:5836091.
Sequence submitted by:
Berkeley Drosophila Genome Project

Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email to bdgpe@fruitfly.berkeley.edu.

FEATURES

Location/Qualifiers

source

1..159455
/organism="Drosophila melanogaster"

/strain="y; cn bw sp"

/db_xref="taxon:7227"

/chromosome="2R"

/map="60B-60C"

/clone="BACR04P18 (D956)"

/clone.lib="RPCI-98 (Roswell Park Cancer Institute

Drosophila melanogaster BAC library, partial ECORI in

pBACe3.6)"

BASE COUNT 41175 a 37532 c 37124 g 43624 t

ORIGIN

Query Match 18.8%; Score 51.4; DB 3; Length 159455;

Best Local Similarity 57.6%; Pred. No. 1.7;

Matches 129; Conservative 0; Mismatches 89; Indels 6; Gaps 2;

Qy 27 ctctactgaactggtcctggatgaacctgtagtcaacgacctgcgcgcatggtcgac 86

Db 29353 CGCTTCTCGGATTCAGCTGGCAGAGATTCGCATCTACGACCTGCCGCCATGTGGAT 29294

Qy 87 ttctg---tcgtcaaacagacgcgcgcagagcctcactacgtcgacacatccatggggacg 143

Db 29293 TACGTCTGGCAGGACGAATCGCCGACGTGCTACTAGTCTGGCCTCGCAGGCACCC 29234

Qy 144 ctggtggtgcgcgcgcgcctcttcgagggccgggtggtgagccagctgaatcc---gcg 200

Db 29233 ACCGTTCTCTGCTGCTCGCAGCGCGGAGTACAATGCGCGTTCGCCAATGCC 29174

Qy 201 gcgctgctcagcgcgcgtggtgctacctcgcacatnaacnccc 244

Db 29173 GCCCTGTTGGCTCCGGTGGCATTTCTCCAGCACCTGAGCAGTCC 29130

RESULT 12

AC012167/c

LOCUS

AC012167 164443 bp DNA linear INV 14-MAR-2001

DEFINITION Drosophila melanogaster, chromosome 2R, region 60B-60C, BAC clone

BACR27M17, complete sequence.

AC012167

AC012167.5 GI:13324743

HTG

SOURCE

fruit fly.

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 164443)

REFERENCE

AUTHORS

Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
Holt,R.A., Evans,C.A., Gocayne,J.D., Ananides,P.G., Brandon,R.C.,
Rogers,Y., An,H., Baldwin,D., Bazon,J., Beeson,K.Y., Busam,D.A.,
Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
Ferreira,S., Frishe,B., Galle,R.F., Garg,N.S., George,R.A.,
Gonzalez,M., Houch,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
Ibegwam,C., Jalili,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
Stapleton,M., Strong,R., Svirskaas,R., Tector,C., Williams,S.M.,
Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
Sequencing of Drosophila chromosome 2R, region 60B-60C
Unpublished

TITLE

JOURNAL

REFERENCE

AUTHORS

2 (bases 1 to 164443)
Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomoton,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
Svirskaas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.

Direct Submission

TITLE

JOURNAL

Submitted (21-OCT-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA

On Mar 14, 2001 this sequence version replaced gi:6838820.

COMMENT

Sequence submitted by:

Berkeley Drosophila Genome Project

Lawrence Berkeley National Laboratory, MS 64-121

Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome
shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive web site (<http://www.fruitfly.org/sequence/>) or send email
to bdgpe@fruitfly.berkeley.edu.

FEATURES

source

1..164443
/organism="Drosophila melanogaster"

/strain="y; cn bw sp"

/db_xref="taxon:7227"

/chromosome="2R"

/map="60B-60C"

/clone="BACR27M17 (D1146)"

/clone.lib="RPCI-98 (Roswell Park Cancer Institute

Drosophila melanogaster BAC library, partial ECORI in

pBACe3.6)"

BASE COUNT 44112 a 38308 c 37703 g 44320 t

ORIGIN

Query Match 18.8%; Score 51.4; DB 3; Length 164443;

Best Local Similarity 57.6%; Pred. No. 1.7;

Matches 129; Conservative 0; Mismatches 89; Indels 6; Gaps 2;

Qy 27 ctctactggaactggtcctggatgaacctgtagtcaacgacctgcgcgcatggtcgac 86

Db 149958 CGCTTCTCGGATTCAGCTGGCAGAGATTCGCATCTACGACCTGCCGCCATGTGGAT 149899

Qy 87 ttctg---tcgtcaaacagacgcgcgcagagcctcactacgtcgacacatccatggggacg 143

Db 149898 TACGTCTGGCAGGACGAATCGCCGACGTGCTACTAGTCTGGCCTCGCAGGCACCC 149839

Qy 144 ctggtggtgcgcgcgcgcctcttcgagggccgggtggtgagccagctgaatcc---gcg 200

Db 149838 ACCGTTCTCTGCTGCTCGCAGCGCGGAGTACAATGCGCGTTCGCCAATGCC 149779

Qy 201 gcgctgctcagcgcgcgtggtgctacctcgcacatnaacnccc 244

Db 149778 GCCCTGTTGGCTCCGGTGGCATTTCTCCAGCACCTGAGCAGTCC 149735

RESULT 13

AE003463/c

LOCUS

AE003463 299975 bp DNA linear INV 04-OCT-2000
DEFINITION Drosophila melanogaster genomic scaffold 142000013386038 section 12
of 15, complete sequence.

ACCESSION

AE003463 AE002575

VERSION

AE003463.1 GI:7291737

KEYWORDS

HTG.

SOURCE

fruit fly.

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE	1 (bases 1 to 299975)	
AUTHORS	Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X., Brannon,R.C., Rogers,Y.H., Blazer,V., Blakes,R.G., Champ,M., Pfeiffer,B.D., Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J., Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y., Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C., Busam,A., Butler,H., Cadieu,E., Center,A., Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferrieria,S., Fleischmann,W., Fostler,C., Gabriellian,A.E., Garg,N.S., Gelbart,W.L., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z., Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J., Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J., Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z., Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A., Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattel,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C., Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nusskern,D.R., Pacleb,J.M., Palazzolo,M., Pittman,G.S., Pan,S., Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K., Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I., Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C., Stapleton,M., Strong,R., Sun,E., Svirskas,R., Tector,C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A., Weinstock,G.M., Weissbach,J., Williams,S.M., Woodage,T., Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F., Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,F.N., Zhong,W., Zhou,X., Zhu,X., Zhu,X., Smith,H.O., Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.	
JOURNAL	The genome sequence of <i>Drosophila melanogaster</i>	
MEDLINE	Science 287 (5461), 2185-2195 (2000)	
REFERENCE	2 (bases 1 to 299975)	
AUTHORS	Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.	
TITLE	Direct Submission	
JOURNAL	Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA	
FEATURES	Location/Qualifiers	
source	1..299975	
	/organism="Drosophila melanogaster"	
	/db_xref="taxon:7227"	
	/chromosome="2R"	
	complement(join(<344..1743,2075..>2486))	
	/gene="CG10339"	
	/product="CT29036"	
	/db_xref="FLYBASE:FBan010339"	
	/db_xref="FLYBASE:FBgn0034972"	
	complement(<344..>2486)	
	/gene="CG10339"	
	/map="60B3-60B5"	
	/db_xref="FLYBASE:FBan010339"	
	/db_xref="FLYBASE:FBgn0034972"	
	complement(join(344..1743,2075..>2486))	
	/gene="CG10339"	
	/note="CG10339 gene product"	
	/codon_start=1	
	/db_xref="FLYBASE:FBan010339"	
	/db_xref="FLYBASE:FBgn0034972"	
	/protein_id="AAF47159.1"	
	/db_xref="GI:7291738"	
	/translation="MDCRANFLAYLAVFLHLVSDYCFVQCVDAPFIHPGNVGIS TYLKMFTQNKAYLGTIRYAHARFPQPDLETFWKGIFNATVPQPCWQNMPSVGK ETEQILKIIQSDSSAQDAERKYEENCLYLVNVPDGLPEINGAYAVVWIIHSGDFSTGS	
mrna		
gene		
CDS		

mrna	PADEVPEQLVFKQKVIIVTFYSRLNIEGFFFTDDGEAGQNYGLMQSAAALVWKKNN FFGDSNTRITLMHGDAVSVALHMTSGEMSKGFGHRAIIMSGNPLSTVKRPYEGS LDOYSSNFACPRRTSMFIOCLRKIDAKRLSENLPQYWSGVPVDFGLSNTSYPIENG PEILFKGSHVRVPIIGVWDMEVILFLKDNLDTDLSPSDLEGFYTDIATVDMHKL V RNWECSNYELISDAINFMVANDSDTDAKKANKHIIISAHTEKFYITPQTFADLIISD SQVSYLFRMRPSVLQDLPSWISPKYFDQIFVWGNPYMTNSVDWMSKTADIIIM TLWANFAKTSNPKTSYVKNAMTNNDSVLLIDESFNTDNLNLRNINRNFWSLYPK ILVFSADCCNSTSGSLVIATSVPICTLSIIIVNF"	
gene	2961..3334	
	/gene="CG13564"	
	/product="CT32944"	
	/db_xref="FLYBASE:FBan0013564"	
	/db_xref="FLYBASE:FBgn0034973"	
	<2961..>3334	
	/gene="CG13564"	
	/map="60B3-60B5"	
	/db_xref="FLYBASE:FBan0013564"	
	/db_xref="FLYBASE:FBgn0034973"	
	3033..3278	
	/gene="CG13564"	
	/note="CG13564 gene product"	
	/codon_start=1	
	/db_xref="FLYBASE:FBan0013564"	
	/db_xref="FLYBASE:FBgn0034973"	
	/protein_id="AAF47160.1"	
	/db_xref="GI:7291739"	
	/translation="MSHEPKPAKSAVKRPVSEETRTDCKLADFLCKGGSGLIIG SAVTLFTTRQTYPIWLGIVGMGVAYDCCOARNQQ"	
	join((3907..4048,4991..5079,5481..5594,5662..6295, 6349..6829,6942..>8012)	
	/gene="CG16786"	
	/product="CT32945"	
	/db_xref="FLYBASE:FBan0016786"	
	/db_xref="FLYBASE:FBgn0034974"	
	<3907..>8012	
	/gene="CG16786"	
	/map="60B3-60B5"	
	/db_xref="FLYBASE:FBan0016786"	
	/db_xref="FLYBASE:FBgn0034974"	
	join(5040..5079,5481..5594,5662..6295,6349..6829, 6942..8012)	
	/gene="CG16786"	
	/note="CG16786 gene product"	
	/codon_start=1	
	/db_xref="FLYBASE:FBan0016786"	
	/db_xref="FLYBASE:FBgn0034974"	
	/protein_id="AAF47161.1"	
	/db_xref="GI:7291740"	
	/translation="MLRYLLILCALGSATCGDVLLTQTVYGVFLDFTTTTIGNTVMVFS PQSSPAFELKSNDTHEIINIETKPKVTKKENVGIKPTPLIVTHSNEIPETETSQI ESNISLNSPIQTLNTPEDLLSRQEQFAEETRYLVNFKSKSEADQRQIYPSKKE PVHGQATRSTEPSIVTATKDSGKFSKVSQVHAVSVKSKPQSRNKPVPISQ PSRTLKSSIQTLQASKPPKSHQSGDRNGSSGTRKPSRPKGRNKHQTSKPTVA PSATETTSRHSYRSKANANAVEPVSVPVSNAPKLNRRPQWQYKSKPKVNIKRN ANPLPTTVQNDTAPDEIPVDIITNQAINNRDLSEAGSQNGVANNNDPDKSKFY VQTLNVEISTPQPFIDTYEATIKTFIFQAGVYKTRFTVTSTIEKIQEETHKE YSEDGPLETENILEATASTDKTLTGSVTLKPIYLGEEETPSLETSTIEKIQEETHK LKTOILPVVYAKRNETIQVTLQVTDYTLTITVQTLTISPLSDNPKSNKDFEGSLD EAGSEINLDLEFGDEDTNGFDAKIKPKLVEAKSNVTSPINLLPESISPEPELQNS LITSTRPVKLETIWESHVPLVRGTETIMRTLKSKSVGVEKTEITVDVATISMTPTFS SQYFSPNPNFPNPLPIAPQPFITSTEVOQSMVTSTSSKVLKLTFGARTAYTPTFS TSVPVATVTRLTATIPQCGQSPFNFPFPYNPFAVYG"	
	join(8642..8712,9542..9726,9786..10020,10075..10553, 10614..10888,10952..11277,11566..11715,11776..12059, 12132..14026)	
	/gene="BcDNA:LD21643"	
	/product="CT10841"	
	/db_xref="FLYBASE:FBan0003231"	
	/db_xref="FLYBASE:FBgn0027522"	
	<8642..>14026	
	/gene="BcDNA:LD21643"	
	/note="CG3231"	
	/map="60B4-60B5"	

	Query Match	Best Local Similarity	58.1%;	Score 49.4;	DB 6;	Length 29736;	
	Matches	98;	Conservative	0;	Mismatches	84;	
					Indels	Gaps	
QY	63	aacgacctgcg	cgccatggttcg	acttcgtcg	ctcaaacagac	gcgcgaagcctcactac	122
Db	15717	AACGTGCTCG	CGCAGTTCG	AGTTCGTTCG	CGCGCAGAACT	CAACAGCTGCTCGCC	15776

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:48:47 ; Search time 23.46 seconds
(without alignments)
426.114 Million cell updates/sec

Title: US-09-699-652-12

Perfect score: 468

Sequence: 1 SSCTRFQLYWNSWDDLVV.....PVAYLXHNXPXNFCWPXR 90

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	458	97.9	90	21	AAV32306
2	279	59.6	410	21	AAV32309
3	188	40.2	405	21	AAV32310
4	181	38.7	286	21	AAV32304
5	180	38.5	410	21	AAV32307
6	172	36.8	394	22	AB71702
7	164	35.0	203	22	AB65717
8	158	33.8	398	14	AAV37302
9	156	33.3	398	7	AAV60724
10	156	33.3	398	7	AAV60658
11	156	33.3	398	17	AAV09383

12	156	33.3	398	22	AAV66086	Human lipase
13	154	32.9	379	15	AAV56870	Canine gastric lip
14	154	32.9	379	17	AAV09382	Dog gastric lipase
15	154	32.9	380	15	AAV56871	Canine gastric lip
16	149	31.8	398	22	AAV63011	Drosophila melanog
17	145	31.0	233	22	AAV61608	Human protein Hp03
18	145	31.0	399	22	AAV67513	Amino acid sequenc
19	145	31.0	656	22	ABG26839	Novel human diagno
20	144	30.8	395	6	AAV50322	Rat lingual lipase
21	141	30.1	399	22	ABV71442	Drosophila melanog
22	141	30.1	435	22	ABV71506	Drosophila melanog
23	138	29.5	221	22	AAV66068	Human TANGO 294 ex
24	138	29.5	390	22	AAV66067	Human TANGO 294 ma
25	138	29.5	409	22	AAV11931	Human CG162 (or C5
26	138	29.5	423	22	AAV66065	Human TANGO 294..
27	138	29.5	457	22	ABV63187	Drosophila melanog
28	136	29.1	399	22	AAV90783	Human shear stress
29	136	29.1	399	22	AAV66061	Human lysosomal ac
30	130	27.8	456	22	ABV68957	Drosophila melanog
31	126	26.9	449	22	AAU30498	Novel human secret
32	122	26.1	416	22	ABV59328	Drosophila melanog
33	121	25.9	434	22	ABV62081	Drosophila melanog
34	119	25.4	341	22	ABV65835	Drosophila melanog
35	117.5	25.1	838	22	ABV61967	Drosophila melanog
36	116	24.8	319	22	ABV65831	Drosophila melanog
37	116	24.8	319	22	ABV62219	Drosophila melanog
38	116	24.8	333	22	ABV65839	Drosophila melanog
39	116	24.8	342	22	ABV60053	Drosophila melanog
40	115	24.6	439	22	ABV63143	Drosophila melanog
41	114	24.4	406	22	ABV66863	Drosophila melanog
42	94	20.1	318	22	ABV66166	Drosophila melanog
43	93	19.9	355	22	ABV66169	Drosophila melanog
44	93	19.9	356	22	ABV66848	Drosophila melanog
45	70.5	15.1	308	21	AAV48793	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAV32306
ID AAV32306 standard; Protein; 90 AA.
XX
AC AAV32306;
XX
DT 28-FEB-2000 (first entry)
XX
XX Corn acid triacylglycerol lipase fragment.
DE DE Triacylglycerol lipase; corn; maize; fatty acid; seed oil;
KW KW vegetable oil; transgenic plant.
XX
OS Zea mays.
XX
XX Key Location/Qualifiers
FH Misc-difference 2 /note= "encoded by TCN"
FT
FT Misc-difference 6 /note= "encoded by CGN"
FT
FT Misc-difference 76 /note= "encoded by GCN"
FT
FT Misc-difference 78 /note= "encoded by ATN"
FT
FT Misc-difference 80 /note= "encoded by ANC"
FT
FT Misc-difference 83 /note= "encoded by TNG"
FT
FT Misc-difference 89 /note= "encoded by ANG"
XX
PN WO9955883-A2.
XX
PD 04-NOV-1999.


```

XX 29-APR-1999; 99WO-US09280.
XX 30-APR-1998; 98US-0083688.
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX Cahoon EB, Cahoon RE, Kinney AJ, Rafalski JA;
XX WPI: 2000-062036/05.
XX N-PSDB; AAZ34959.
XX Novel plant triacylglycerol lipase polynucleotides used to alter the
XX level of the enzyme in transgenic plants -
XX Claim 5; Page 50-51; 65pp; English.
XX This sequence represents the entire soybean acid triacylglycerol
XX lipase (TAGL), as deduced from the nucleotide sequence of isolated
XX cDNA clones (see AAZ34959). Novel acid and neutral TAGL polypeptides
XX (see AAY32301-17) and polynucleotides (see AAZ34950-66) from corn,
XX Catalpa, rice, soybean and wheat tissues are disclosed. The enzymes
XX may be prepared recombinantly and used to raise antibodies, which
XX are used for detecting the enzymes in situ in cells or in vitro in
XX cell extracts. The polynucleotides may be used to create transgenic
XX plants in which the TAGL levels are present at higher or lower levels
XX than normal, or in cell types or developmental processes where they are
XX not normally found. This would alter the level of triacylglycerol and
XX cholesteryl esters found in those cells. Accumulation of fatty acids
XX with unusual structures may be a positive phenotype in plants used for
XX foods. In addition, it may be desirable to eliminate expression of TAGL
XX genes for certain applications. TAGL enzymes may also be useful for the
XX processing of plant seed oils and for the development of novel seed
XX oils. The TAGL enzymes can also be used as targets to facilitate the
XX design and/or identification of inhibitors of those enzymes that may be
XX useful as herbicides. This is desirable because inhibition of the
XX activity of either of the enzymes could lead to an inhibition of plant
XX growth.
XX Sequence 405 AA;

Query Match 40.2%; Score 188; DB 21; Length 405;
Best Local Similarity 40.3%; Pred. No. 5.3e-16;
Matches 29; Conservative 23; Mismatches 20; Indels 0; Gaps 0;

QY 10 YNWSWDDLVDLPAMVDVVKQGRPHYVGHSGMTGLVALAASFSEGRVVSQKSAALL 69
DB 144 fvdswqelalydvaeminyinsvtnskifvghsggtiislaaftgpeivekveaaall 203

QY 70 TPVAYLXHXXNP 81
DB 204 spisyldhvsap 215

RESULT 4
AAY32304
ID AAY32304 standard; Protein; 286 AA.
XX
XX AC AAY32304;
XX
XX DT 28-FEB-2000 (first entry)
XX
XX DE Corn acid triacylglycerol lipase.
XX
XX KW Triacylglycerol lipase; corn; maize; fatty acid; seed oil;
XX
XX KW vegetable oil; transgenic plant.
XX
XX OS Zea mays.
XX
XX FH Key Location/Qualifiers
XX
XX FT Misc-difference 16
XX /note= "encoded by NAC"

```

```

XX WO9955883-A2.
XX PN
XX PD 04-NOV-1999.
XX
XX 29-APR-1999; 99WO-US09280.
XX PF
XX 30-APR-1998; 98US-0083688.
XX PR
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX PA
XX Cahoon EB, Cahoon RE, Kinney AJ, Rafalski JA;
XX WPI: 2000-062036/05.
XX DR N-PSDB; AAZ34953.
XX
XX Novel plant triacylglycerol lipase polynucleotides used to alter the
XX level of the enzyme in transgenic plants -
XX Claim 5; Page 40-41; 65pp; English.
XX This sequence represents a substantial portion of corn acid
XX triacylglycerol lipase (TAGL), as deduced from the nucleotide
XX sequence of a contig (see AAZ34953) assembled from isolated leaf and
XX shoot cDNA clones. Novel acid and neutral TAGL polypeptides (see
XX AAY32301-17) and polynucleotides (see AAZ34950-66) from corn, Catalpa,
XX rice, soybean and wheat tissues are disclosed. The enzymes
XX may be prepared recombinantly and used to raise antibodies, which
XX are used for detecting the enzymes in situ in cells or in vitro in
XX cell extracts. The polynucleotides may be used to create transgenic
XX plants in which the TAGL levels are present at higher or lower levels
XX than normal, or in cell types or developmental processes where they are
XX not normally found. This would alter the level of triacylglycerol and
XX cholesteryl esters found in those cells. Accumulation of fatty acids
XX with unusual structures may be a positive phenotype in plants used for
XX foods. In addition, it may be desirable to eliminate expression of TAGL
XX genes for certain applications. TAGL enzymes may also be useful for the
XX processing of plant seed oils and for the development of novel seed
XX oils. The TAGL enzymes can also be used as targets to facilitate the
XX design and/or identification of inhibitors of those enzymes that may be
XX useful as herbicides. This is desirable because inhibition of the
XX activity of either of the enzymes could lead to an inhibition of plant
XX growth.
XX Sequence 286 AA;

Query Match 38.7%; Score 181; DB 21; Length 286;
Best Local Similarity 44.2%; Pred. No. 2.8e-15;
Matches 34; Conservative 15; Mismatches 28; Indels 0; Gaps 0;

QY 1 SSCTLRFOLYWNWSWDDLVDLPAMVDVVKQGRPHYVGHSGMTGLVALAASFSEGRV 60
DB 146 stlsvhdklfvdswqgdlaeydvlamlsyvtvaqskilyvghsggtimglaaftmpetv 205

QY 61 SOLKSAALLTPVAYLXH 77
DB 206 kmisaallcpisylh 222

RESULT 5
AAY32307
ID AAY32307 standard; Protein; 410 AA.
XX
XX AC AAY32307;
XX
XX DT 28-FEB-2000 (first entry)
XX
XX DE Rice acid triacylglycerol lipase.
XX
XX KW Triacylglycerol lipase; rice; fatty acid; seed oil;
XX
XX KW vegetable oil; transgenic plant.
XX
XX

```


XX PS Example; Fig 5; 60pp; English.

XX CC The inventors claim a DNA sequence encoding AAP60723 linked to a

CC promoter upstream and a gene for a polypeptide downstream.

CC Particular examples are the yeast enzyme AMG, the mammalian enzyme,

CC gastric lipase and the mammalian lymphokine, interferon-alpha2.

XX SQ Sequence 398 AA;

Query Match 33.3%; Score 156; DB 7; Length 398;

Best Local Similarity 47.0%; Pred. No. 8e-12;

Matches 31; Conservative 13; Mismatches 20; Indels 2; Gaps 2;

QY 10 YNWSWDDLVDLPAMVDVVKOTGQKP-HYVGHSMGTLVALAAAFSEG-RVVSOLKSAA 67

Db 137 fwafsdemakydldpatidflvkkctgkqlhyvghsqgttigfalfstnpslaktikfy 196

QY 68 LLTPVA 73

Db 197 alapva 202

RESULT 10

ID AAP60658 standard; Protein; 398 AA.

XX AAP60658;

XX 22-AUG-1991 (first entry)

DE Sequence of human pregastric lipase.

KW Cystic fibrosis therapy; enzyme; lipase deficiency.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..19

FT Protein 20..398

XX WO8601532-A.

XX 13-MAR-1986.

XX 15-AUG-1985; 85WO-GB00364.

XX 21-AUG-1984; 84GB-0021210.

PR 15-AUG-1985; 85WO-GB00364.

PR 01-JAN-1986; 86GB-0008897.

XX (CELL-) CELLTech LTD.

PA (LOWE/) LOWE P A.

XX Lowe PA;

XX WPI; 1986-081634/12.

DR N-PSDB; AAN60566.

XX New gastric lipase protein, esp. of human origin - for treating

PT lipase deficiency, and DNA sequences coding for it

XX Disclosure; Fig 3; 39pp; English.

XX The inventors claim a pregastric lipase protein and a gene encoding

CC it. Gastric lipase protein is useful for oral administration to

CC treat lipase deficiency, e.g. cystic fibrosis or pancreatitis.

XX Sequence 398 AA;

Query Match 33.3%; Score 156; DB 7; Length 398;

Best Local Similarity 47.0%; Pred. No. 8e-12;

Matches 31; Conservative 13; Mismatches 20; Indels 2; Gaps 2;

QY 10 YNWSWDDLVDLPAMVDVVKOTGQKP-HYVGHSMGTLVALAAAFSEG-RVVSOLKSAA 67

Db 137 fwafsdemakydldpatidflvkkctgkqlhyvghsqgttigfalfstnpslaktikfy 196

QY 68 LLTPVA 73

Db 197 alapva 202

RESULT 11

ID AAW09383 standard; Protein; 398 AA.

XX AAW09383;

XX 19-AUG-1997 (first entry)

DE Human gastric lipase protein sequence.

XX Duodenal; gastric; lipase; transgenic; plant; recombinant; extract;

KW food; absorption; fat; pancreatic; cystic fibrosis; exocrine; dairy;

KW hydrolysis; trans-esterification; substrate; enzyme; biofuel; human.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..19

FT Protein 20..999

FT /note= "mature protein"

XX WO9633277-A2.

XX 24-OCT-1996.

XX 19-APR-1996; 96WO-FR00606.

PR 20-APR-1995; 95FR-0004754.

XX (BIOC-) BIOCEM SA.

PA (LJO) INST RECH JOUVEINAL.

XX Baudino S, Benicourt C, Cudrey C, Gruber V, Lenee P;

PI Merot B;

XX WPI; 1996-485783/48.

DR N-PSDB; AAT58916.

XX Recombinant human or canine pre-duodenal lipase prodn. in transgenic

PT plants - useful for facilitating absorption of fat, as bio-catalysts

PT and for prodn. of bio-fuel

XX Claim 3; Fig 5; 130pp; French.

XX This is the amino acid sequence of the human pre-duodenal (i.e. gastric)

CC lipase enzyme. The sequence can be used to generate transgenic plants

CC producing recombinant lipase in an enzymatically active form.

CC Alternatively amino acids 20-23 or 20-73 (1-4 or 1-54 of the mature

CC protein, respectively) can be deleted to form the derivatives designated

CC delta-4 or delta-54 respectively. Plants, or their extracts, expressing

CC the lipases or the truncated derivatives, can be used:

CC (a) as pharmaceuticals or food to facilitate absorption of fat, either

CC in healthy subjects or in patients with inadequate levels of

CC gastric/pancreatic lipase e.g. those with cystic fibrosis or exocrine

CC pancreatic insufficiency, in the elderly or in patients undergoing medical

CC treatment which alters fat adsorption; (b) for performing industrial or

CC agricultural reactions, e.g. in processing of fats or in the dairy

CC industry, for hydrolysis or transesterification reactions, etc., where

CC the plant material may provide both enzyme and substrate. The transgenic

CC plants can also be used for biofuel production.

	Query Match	33.3%; Score 156; DB 22; Length 398;
	Best Local Similarity	47.0%; Pred. No. 8e-12;
	Matches	31; Conservative 13; Mismatches 20; Indels 2; Gaps
QY	10 YWNKSWDDLVDLVPAMVDFVVKQTGKP-HYVGHSKMTLVLAARFSEG-RVVSQLKSAA	67
Dd	137 fwafsfdemakydipatidfvkktgqqlhyvghsggttgfiastnpslakrikty	179
QY	68 LLTPVA 73	
Dd	197 alapva 202	
 RESULT 13		
AAR56870		
ID	AAR56870 standard; Protein; 379 AA.	
XX	AAR56870:	
XX	20-FEB-1995 (first entry)	
XX	Canine gastric lipase.	
DE		
XX	Canine gastric lipase; CGL; dog; fat; mucoviscidosis; enzyme;	
KW	bioconversion; exocrine pancreatic insufficiency.	
XX	Canis familiaris.	
XX	WO9413816-A.	
XX	23-JUN-1994.	
XX	16-DEC-1993; 93WO-FR01260.	
XX	16-DEC-1992; 92FR-0015201.	
PR	(LJOU) INST RECH JOUVEINAL.	
PA	Benicourt C, Blanchard C, Junien J;	
PI	WPI: 1994-217890/26.	
DR	N-PSDB; AAQ68388.	
DR		
XX	Recombinant canine gastric lipase and nucleic acid encoding it -	
PT	are used for improving absorption of ingested fat, treating	
PT	mucoviscidosis etc. and in enzymatic bio-conversions	
XX	Claim 13; Fig 9A; 52pp; French.	
PS	The sequence given below is the sequence of figure 9A, altered	
CC	according to the amendments described on page 2 of the appended	
CC	letter.	
CC	CGL is used to improve absorption of ingested fat, in healthy and	
CC	sick patients (e.g. having altered levels of gastric lipase); to	
CC	treat conditions associated with insufficiency (or lack) of lipases,	
CC	esp. mucoviscidosis or exocrine pancreatic insufficiency and partic.	
CC	where immobilised, for bioconversions, e.g. hydrolysis or	
CC	transesterification (other mammalian gastric lipases, or derivs.,	
CC	can be used in this application).	
XX		
SQ	Sequence 379 AA;	
 Query Match 32.9%; Score 154; DB 15; Length 379;		
	Best Local Similarity	45.5%; Pred. No. 1.4e-11;
	Matches	30; Conservative 15; Mismatches 19; Indels 2; Gaps
QY	10 YWNKSWDDLVDLVPAMVDFVVKQTGO-KPHYVGHSMGTLVLAARFSEG-RVVSQLKSAA	67
Dd	118 fwafsfdemakydipatidfkktgqdklhyvghsggttgfiastnpslakrikty	177

OY 68 LITPVA 73
 Db 178 alapva 183

RESULT 14
 AAW09382
 ID AAW09382 standard; Protein; 379 AA.

XX AC AAW09382;
 XX DT 19-AUG-1997 (first entry)

XX DE Dog gastric lipase protein sequence.

XX KW Dog; duodenal; gastric; lipase; transgenic; plant; recombinant; extract;
 KW food; absorption; fat; pancreatic; cystic fibrosis; exocrine; dairy;
 KW hydrolysis; trans-esterification; substrate; enzyme; biofuel.

XX OS Canis familiaris.

XX PN W09633277-A2.

XX PD 24-OCT-1996.

XX PF 19-APR-1996; 96WO-FR00606.

XX PR 20-APR-1995; 95FR-0004754.

XX PA (BIOC-) BIOCEM SA.

XX PA (LJOU) INST RECH JOUVEINAL.

XX PI Baudino S, Benicourt C, Cudrey C, Gruber V, Lenee P;

XX PI Merot B;

XX XX WPI; 1996-485783/48.

DR N-PSDB; AAT58915.

XX Recombinant human or canine pre-duodenal lipase prodn. in transgenic
 PT plants - useful for facilitating absorption of fat, as bio-catalysts
 PT and for prodn. of bio-fuel

XX Claim 2; Fig 2; 130pp; French.

XX This is the amino acid sequence of the dog pre-duodenal (i.e. gastric)
 CC lipase enzyme. The sequence can be used to generate transgenic plants
 CC producing recombinant lipase in an enzymatically active form.
 CC Alternatively the first 4 or 54 amino acids of this sequence can be
 CC deleted to form the derivatives designated delta-4 or delta-54
 CC respectively. Plants, or their extracts, expressing the lipases or the
 CC truncated derivatives, can be used: (a) as pharmaceuticals or food to
 CC facilitate absorption of fat, either in healthy subjects or in patients
 CC with inadequate levels of gastric/pancreatic lipase e.g. those with
 CC cystic fibrosis or exocrine pancreatic insufficiency, in the elderly or in
 CC treatment which alters fat adsorption; (b) for performing industrial or
 CC agricultural reactions, e.g. in processing of fats or in the dairy
 CC industry, for hydrolysis or transesterification reactions, etc., where
 CC the plant material may provide both enzyme and substrate. The transgenic
 CC plants can also be used for biofuel production.

SQ Sequence 379 AA;

Query Match 32.9%; Score 154; DB 17; Length 379;
 Best Local Similarity 45.5%; Pred. No. 1.4e-11;
 Matches 30; Conservative 15; Mismatches 19; Indels 2; Gaps 2;

OY 10 YWNWSDLLVNDLPAMVDVVKQTGQ-KPHYVGHSMGTLVALAAAFSEG-RVVSOLKSA 67

Db 118 fwafsdemakydldpatidflkktgqdklhvghsggttgifafstnpklakrikf 177

OY 68 LITPVA 73

Db 178 alapva 183

Db 178 alapva 183

RESULT 15

AAR56871
 ID AAR56871 standard; Protein; 380 AA.

XX AC AAR56871;

XX DT 20-FEB-1995 (first entry)

XX DE Canine gastric lipase (Met-CGL).

XX KW Canine gastric lipase; CGL; dog; fat; mucoviscidosis; enzyme;
 KW bioconversion; exocrine pancreatic insufficiency.

XX OS Canis familiaris.

XX PN W09413816-A.

XX PD 23-JUN-1994.

XX PF 16-DEC-1993; 93WO-FR01260.

XX PR 16-DEC-1992; 92FR-0015201.

XX PA (LJOU) INST RECH JOUVEINAL.

XX PI Benicourt C, Blanchard C, Junien J;

XX DR WPI; 1994-217890/26.

DR N-PSDB; AAQ68389.

XX Recombinant canine gastric lipase and nucleic acid encoding it -
 PT are used for improving absorption of ingested fat, treating
 PT mucoviscidosis etc. and in enzymatic bio-conversions

XX Claim 14; Fig 9A; 52pp; French.

XX The sequence given below is the sequence of figure 9A, altered
 CC according to the amendments described on page 2 of the appended
 CC letter.
 CC CGL is used to improve absorption of ingested fat, in healthy and
 CC sick patients (e.g. having altered levels of gastric lipase); to
 CC treat conditions associated with insufficiency (or lack) of lipases,
 CC esp. mucoviscidosis or exocrine pancreatic insufficiency and partic.
 CC where immobilised, for bioconversions, e.g. hydrolysis or
 CC transesterification (other mammalian gastric lipases, or derivs.,
 CC can be used in this application).

XX SQ Sequence 380 AA;

Query Match 32.9%; Score 154; DB 15; Length 380;
 Best Local Similarity 45.5%; Pred. No. 1.4e-11;
 Matches 30; Conservative 15; Mismatches 19; Indels 2; Gaps 2;

OY 10 YWNWSDLLVNDLPAMVDVVKQTGQ-KPHYVGHSMGTLVALAAAFSEG-RVVSOLKSA 67

Db 119 fwafsdemakydldpatidflkktgqdklhvghsggttgifafstnpklakrikf 178

OY 68 LITPVA 73

Db 179 alapva 184

Search completed: October 10, 2002, 02:10:27
 Job time: 1300 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:57 ; Search time 21.51 Seconds
(without alignments)
723.828 Million cell updates/sec

Title: US-09-699-652-12
Perfect score: 468
Sequence: 1 SSGTTRFQLYWNKSWDDLTV.....PVAYLXHXNPNKNGPXPXR 90

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL19.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rviris.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	170	36.3	344	10 Q9SKL5	Q9skl5 arabidopsis
2	164	35.0	203	5 Q9W195	Q9w195 drosophila
3	157	33.5	395	11 Q9D6N8	Q9d6n8 mus musculus
4	155	33.1	395	11 Q9D7C5	Q9d7c5 mus musculus
5	155	33.1	395	11 Q9D798	Q9d798 mus musculus
6	155	33.1	395	11 Q9D796	Q9d796 mus musculus
7	155	33.1	395	11 Q9D767	Q9d767 mus musculus
8	155	33.1	395	11 Q9D766	Q9d766 mus musculus
9	155	33.1	395	11 Q9D760	Q9d760 mus musculus
10	155	33.1	395	11 Q9D6X0	Q9d6x0 mus musculus
11	155	33.1	395	11 Q9D6T5	Q9d6t5 mus musculus
12	155	33.1	395	11 Q9D6S5	Q9d6s5 mus musculus
13	155	33.1	395	11 Q9D6Q6	Q9d6q6 mus musculus
14	155	33.1	395	11 Q9D6Q3	Q9d6q3 mus musculus
15	155	33.1	395	11 Q9D6L9	Q9d6l9 mus musculus
16	155	33.1	395	11 Q9D6L1	Q9d6l1 mus musculus

17	155	33.1	395	11 Q9CpP8	Q9cpg8 mus musculus
18	155	33.1	395	11 Q9CpP7	Q9cpg7 mus musculus
19	154	32.9	395	11 Q9B6Q2	Q9b6q2 mus musculus
20	154	32.9	395	11 Q9D6P3	Q9d6p3 mus musculus
21	149	31.8	395	11 Q9D6M9	Q9d6m9 mus musculus
22	149	31.8	398	5 Q9V796	Q9v796 drosophila
23	147.5	31.5	411	5 Q20449	Q20449 caenorhabdi
24	145.5	31.1	404	5 Q16956	Q16956 caenorhabdi
25	145.5	31.1	405	5 Q93789	Q93789 caenorhabdi
26	141	30.1	399	5 Q9VPE9	Q9vpe9 drosophila
27	141	30.1	435	5 Q9VG46	Q9vg46 drosophila
28	141	30.1	616	5 Q77107	Q77107 plodia inte
29	139.5	29.8	351	5 Q95X33	Q95x33 caenorhabdi
30	139.5	29.8	403	5 Q61866	Q61866 caenorhabdi
31	138	29.5	374	4 Q96LG2	Q96lg2 homo sapien
32	138	29.5	457	5 Q9VKT2	Q9vkt2 drosophila
33	136	29.1	399	4 Q16529	Q16529 homo sapien
34	136	29.1	399	4 Q96EJ0	Q96ej0 homo sapien
35	135	28.8	318	10 Q9FMT3	Q9fmt3 arabidopsis
36	130	27.8	456	5 Q9VKS5	Q9vks5 drosophila
37	130	27.8	684	5 Q95XV1	Q95xv1 caenorhabdi
38	129	27.6	397	11 Q9Z0M5	Q9z0m5 mus musculus
39	126	26.9	559	5 Q17219	Q17219 bombyx mori
40	124	26.3	397	11 Q64194	Q64194 rattus sp.
41	123	26.3	426	5 Q17766	Q17766 caenorhabdi
42	122	26.1	411	5 Q94252	Q94252 caenorhabdi
43	122	26.1	416	5 Q9VQ05	Q9vq05 drosophila
44	121	25.9	434	5 Q9VKT9	Q9vkt9 drosophila
45	119	25.4	341	5 Q9VG48	Q9vg48 drosophila

ALIGNMENTS

RESULT 1

ID	Q9SKL5	PRELIMINARY;	PRT;	344 AA.
AC	Q9SKL5;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	POTATIVE LYOSOMAL ACID LIPASE.			
GN	AT2G15230.			
OS	Arabidopsis thaliana (Mouse-ear cross).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.			
OX	NCBI_TaxID=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. COLUMBIA;			
RX	MEDLINE=20083487; PubMed=10617197;			
RA	Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,			
RA	Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,			
RA	Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,			
RA	Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,			
RA	Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,			
RA	Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,			
RA	Salzberg S.L., Fraser C.M., Venter J.C.;			
RT	"Sequence and analysis of chromosome 2 of the plant Arabidopsis			
RL	thaliana.";			
RL	Nature 402:761-768(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. COLUMBIA;			
RA	Lin X.;			
RA	EMBL: AC006298; AAD25569.1;			
DR	InterPro; IPR000379; Est_lip_thloest_actsite.			
DR	InterPro; IPR000734; Lipase.			
DR	PROSITE; PS00120; LIPASE_SER; 1.			
SQ	SEQUENCE 344 AA; 38868 MW; 7BC4ED067E3602C4 CRC64;			

DE 2310051B21RIK PROTEIN.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
ARakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK009433; BAB26283.1; -;
DR MGd; MGI:1914967; 2310051B21RIK.
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR000379; Est_lip_thioest_actsite.
DR InterPro; IPR000734; Lipase.
DR Pfam; PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44603 MW; D3FD8B6FEA671E3E CRC64;

Query Match 33.1%; Score 155; DB 11; Length 395;
Best Local Similarity 41.9%; Pred.No. 7.9e-10;
Matches 31; Conservative 15; Mismatches 26; Indels 2; Gaps

QY 10 YNNWSDDLVDLPAMVDFVVKQTGQ-KPHYVGSHSGMTGLVALAAAFSEG-RVVSOLKSAA 67
: : |:|::: |||||:::||| ||||| |::| : :|
Db 136 FWAFSDEMAYDLPATIDFIVKTGQEKHYHVGSHSGTIGTFIAFNPAALAKKIRFY 195
: : |:|::: |||||:::||| ||||| |::| : :|

QY 68 LLTPVAYLXHNXP 81
: ||| : :
Db 196 ALAPVATVKYTSP 209

RESULT 6
Q9D796 ID Q9D796 PRELIMINARY; PRT; 395 AA.
AC Q9D796;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 2310051B21RIK PROTEIN.
GN 2310051B21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
ARakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK009359; BAB26240.1; -;
DR MGd; MGI:1914967; 2310051B21RIK.
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR000379; Est_lip_thioest_actsite.
DR InterPro; IPR000734; Lipase.
DR Pfam; PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44748 MW; 3ED43992458DE058 CRC64;

Query Match 33.1%; Score 155; DB 11; Length 395;
Best Local Similarity 41.9%; Pred.No. 7.9e-10;
Matches 31; Conservative 15; Mismatches 26; Indels 2; Gaps

QY 10 YNNWSDDLVDLPAMVDFVVKQTGQ-KPHYVGSHSGMTGLVALAAAFSEG-RVVSOLKSAA 67
: : |:|::: |||||:::||| ||||| |::| : :|
Db 136 FWAFSDEMAYDLPATIDFIVKTGQEKHYHVGSHSGTIGTFIAFNPAALAKKIRFY 195
: : |:|::: |||||:::||| ||||| |::| : :|

QY 68 LLTPVAYLXHNXP 81
: ||| : :
Db 196 ALAPVATVKYTSP 209

RESULT 5
Q9D798 ID Q9D798 PRELIMINARY; PRT; 395 AA.
AC Q9D798;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

	Query Match	33.18;	Score 155;	DB 11;	Length 395;
	Best Local Similarity	41.96;	Pred. No. 7.9e-10;		
	Matches	31; Conservative	15; Mismatches	26; Indels	2; Gaps
Qy	10 YWNWSDDLVDNDLPAMVDFVKQTGO-KPHYVGHSMTLVALAAFSEG-RVVSOLKSA	67	:	: : : :	:
Dd	136 FWAFSEDEMAYDLPAIDFIVQKTGQEKIHYVGHSOGTTIGTAFSTNPALAKIRFY	195	:	: : : :	:
Qy	68 LLTPAVYLXHNXP	81		:	
Dd	196 ALAPVATVKYTESP	209		:	
RESULT	9				
ID	Q9D760	PRELIMINARY:	PRT:	395 AA.	
AC	Q9D760;				
DT	01-JUN-2001 (TrEMBLrel. 17, Created)				
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)				
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)				
DE	2310051B21RIK PROTEIN.				
GN	2310051B21RIK.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
NCBI_TaxId	10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C57BL/6J; TISSUE=TONGUE;				
RX	MEDLINE=21085660; PubMed=11217851;				
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,				
RA	Aizawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,				
RA	Aizawa K., Izawa M., Nishi K., Kiyoasa H., Kondo S., Yamanaoka I.,				
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,				
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavanti T.,				
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,				
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,				
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,				
RA	Sakai K., Okido T., Furuno M., Anon H., Baldarelli R., Barsh G.,				
RA	Blake J.J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,				
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,				
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,				
RA	L Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mommaerts P.,				
RA	Nordone P., Ring B., Schenwald M., Rodriguez I., Sakamoto N.,				
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,				
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,				
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,				
RA	Hayashizaki Y.;				
FT	"Functional annotation of a full-length mouse cDNA collection.";				
RL	Nature 405:685-690(2001).				
DR	EBML; AK009560; BAB26359.1; -;				
DR	MG; MG1:1914967; 2310051B21rik.				
DR	InterPro; IPRO000073; Abhydrolase.				
DR	InterPro; InterPro; IPRO000379; Est_lip_thioest_actsite.				
DR	InterPro; IPRO00734; Lipase.				
DR	Pfam; PF00561; abhydrolase. 1.				
DR	PROSITE; PS00120; LIPASE_SER; 1.				
SO	SEQUENCE	395 AA;	44685 MW;	FEF9GB65EA670BEE CRC64;	
Query Match	33.18;	Score 155;	DB 11;	Length 395;	
Best Local Similarity	41.96;	Pred. No. 7.9e-10;			
Matches	31; Conservative	15; Mismatches	26; Indels	2; Gaps	2;
Qy	10 YWNWSDDLVDNDLPAMVDFVKQTGO-KPHYVGHSMTLVALAAFSEG-RVVSOLKSA	67	:	: : : :	:
Dd	136 FWAFSEDEMAYDLPAIDFIVQKTGQEKIHYVGHSOGTTIGTAFSTNPALAKIRFY	195	:	: : : :	:
Qy	68 LLTPAVYLXHNXP	81		:	
Dd	196 ALAPVATVKYTESP	209		:	

[illegible]

RESULT	13
Q9D6Q6	
ID	Q9D6Q6 PRELIMINARY; PRT; 395 AA.
AC	Q9D6Q6;
AD	Q9D6Q6;
DT	01-JUN-2001 (Tremblrel. 17, Created)
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE	ADULT MALE TONGUE CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE	CLONE: J2310067K20, FULL INSERT SEQUENCE.
GN	Z310051B2IRIK.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6J; TISSUE=TONGUE;
RX	MEDLINE=21085660; Pubmed=11217851;
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Iishi Y.,
RA	Arakawa T., Hara A., Fushimi Y., Konno H., Adachi J., Fukuda S.,
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA	Fleischmann W., Gaasterland T., Glissi C., King B., Kochava H.,
RA	Kuehl P., Lewis S., Matsuo Y., Nikolaio I., Pesole G., Quackenbush
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio
RA	Sakai K., Okido T., Furuno M., Anon H., Baldarelli R., Barsh G.,
RA	Blake J., Boirelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA	Gustingich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA	Hayashizaki Y.;

Search completed: October 10, 2002, 02:13:37
Job time: 220 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: October 10, 2002, 02:10:57 ; Search time 12.42 Seconds
(without alignments)
280.576 Million cell updates/sec

Title: US-09-699-652-12

Perfect score: 468

Sequence: 1 SSCITRFQLYNWSWDLV.....PVAYLHXNXPNNXPGWPXR 90

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB ID	Description
1	172	36.8	394	1	LIP3_DROME	O46108 drosophila
2	156	33.3	398	1	LIPG_HUMAN	P07098 homo sapien
3	154	32.9	398	1	LIPG_CANFA	P80035 canis famli
4	152	32.5	395	1	LIPG_RAT	P04634 rattus norv
5	145	31.0	397	1	LIPG_BOVIN	Q29458 bos taurus
6	136	29.1	399	1	LICH_HUMAN	P38571 homo sapien
7	129	27.6	397	1	LICH_MOUSE	Q920m5 mus musculu
8	124	26.5	397	1	LICH_RAT	Q64194 rattus norv
9	115	24.6	439	1	LIP1_DROME	O46107 drosophila
10	73	15.6	604	1	DNAK_BACME	P05646 bacillus me
11	71.5	15.3	373	1	ACOC_ALCEU	P27747 alcaligenes
12	69	14.7	548	1	TGL1_YEAST	P34163 saccharomyc
13	68	14.5	799	1	RS64_MOUSE	Q92268 mus musculu
14	67.5	14.4	204	1	RS5_HUMAN	P46782 homo sapien
15	65	13.9	301	1	MPT5_MYCLE	Q05868 mycobacteri
16	65	13.9	583	1	PHBC_AZOCA	O66392 a poly-beta
17	64.5	13.8	804	1	RS64_HUMAN	O95294 homo sapien
18	64	13.7	204	1	RS5_MOUSE	P97461 mus musculu
19	64	13.7	204	1	RS5_RAT	P24050 rattus norv
20	64	13.7	228	1	RS5_DROME	Q24186 drosophila
21	63.5	13.6	337	1	RFAL_SALTY	P19816 salmonella
22	63	13.5	396	1	DXR_PSEAE	Q9Kgu6 pseudomonas
23	63	13.5	616	1	TYPA_HAEIN	P44910 haemodionas
24	62	13.2	1450	1	DXR_HELPJ	Q92ml6 helicobacte
25	62	13.2	1450	1	CFTR_RABIT	Q00554 oryctolagus
26	61.5	13.1	254	1	YBFF_ECOLI	P75736 escherichia
27	61.5	13.1	268	1	ESL2_MYCPN	P75311 mycoplasma
28	61.5	13.1	305	1	MDLA_PENCA	P25234 penicillium
29	61.5	13.1	1419	1	LYS2_SCHPO	P40976 schizosacch
30	61	13.0	610	1	DNAK_BACSU	P17820 bacillus su
31	61	13.0	612	1	DNAK_LISMO	Q95544 listeria mo
32	61	13.0	612	1	H570_LISIN	Q92bn8 listeria in
33	60.5	12.9	197	1	RS5_CICAR	O65731 cicer ariet

RESULT 1

ID	LIP3_DROME	STANDARD;	PRT;	394 AA.
AC	O46108;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Lipase 3 precursor (EC 3.1.1.-)			
GN	LIP3 OR CG8823.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CANTON-S;			
RX	MEDLINE=98227315; PubMed=9566193;			
RA	Pistillo D., Manzi A., Tino A., Pilo Boyl P., Graziani F., Malva C.;			
RT	"The Drosophila melanogaster lipase homologs: a gene family with			
RT	tissue and developmental specific expression."			
RL	J. Mol. Biol. 276:877-885(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BERKELEY;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers J.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
RA	Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Delcher A., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Duan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,			
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,			
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
RA	Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,			
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleeb J.M.,			
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,			
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,			
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,			
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,			

ALIGNMENTS

34	60.5	12.9	275	1	YAIM_HAEIN	P44556 haemophilus
35	60.5	12.9	375	1	ADH3_YEAST	P07246 saccharomyc
36	60.5	12.9	424	1	YIAN_ECOLI	P37675 escherichia
37	60	12.8	402	1	YD29_SCHPO	O13716 schizosacch
38	60	12.8	473	1	LIP1_RAT	P54316 rattus norv
39	60	12.8	606	1	DNAK_BACTR	Q9Kws7 bacillus th
40	60	12.8	607	1	DNAK_BACST	O45551 bacillus st
41	59.5	12.7	83	1	RSS_ARATH	P51427 arabidopsis
42	59.5	12.7	173	1	GRAM_MOUSE	P51437 mus musculu
43	59	12.6	349	1	YGB0_ECOLI	Q57261 escherichia
44	59	12.6	398	1	DXR_ECOLI	P45568 escherichia
45	59	12.6	590	1	DNAK_STRMU	O06942 streptococc

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -!- TISSUE SPECIFICITY: FAT BODY.
 CC -!- DEVELOPMENTAL STAGE: ONLY AT LARVAL STAGES.
 CC -!- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
 CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.). ALSO SIMILAR TO
 CC LEPIDOPTERAN EGG-SPECIFIC AND YOLK PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Y14367; CAA74737.1; -
 CC EMBL; AE003699; AAF34935.1; -
 CC FlyBase; FBgn0023495; Lip3.
 CC InterPro; IPR000073; Abhydrolase.
 CC InterPro; IPR000379; Est_lip_thioest_actsite.
 CC InterPro; IPR000734; Lipase.
 CC Pfam; PF00561; abhydrolase.1.
 CC PROSITE; PS00120; Lipase_SER; 1.
 CC Hydrolyase; Lipid degradation; Signal; Glycoprotein.
 KW SIGNAL
 FT CHAIN 1 20 POTENTIAL.
 FT ACT_SITE 164 164 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 369 369 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CARBOHYD 131 131 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT SEQUENCE 394 AA; 44901 MW; A718D1D743673802 CRC64;
 SQ
 Query Match 36.8%; Score 172; DB 1; Length 394;
 Best Local Similarity 44.2%; Pred. No. 2.5e-12;
 Matches 34; Conservative 18; Mismatches 23; Indels 2; Gaps 2;
 QY 7 FOLYNNWSDDLVLVNDLPAMVDVVKQTGQKP-HYVGHSMGTVALAAAFSE-GRVVSQLK 64
 Db :||:||||: :||:||||: :||:||||: ||||| ||||| ||||| :||: |||||
 126 WQIFWNSWNEIGMDVPMIDYVLAKTQVQVGHSGQTTVYLVMSRPEYNDKIK 185
 QY 65 SAALLTPVAYLXHXNP 81
 Db ||||| ||||| :||: ||||| :||: ||||| ||||| ||||| ||||| :||: |||||
 186 SAHLGLPAAYMGNMKSP 202
 QY 68 LLTPVA 73
 Db |||||
 197 ALAPVA 202
 RESULT 2
 LIPG_HUMAN STANDARD; PRT; 398 AA.
 AC P07098;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Triacylglycerol lipase, gastric precursor (EC 3.1.1.3) (Gastric
 DE lipase) (GL).
 GN LIFF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=87299724; PubMed=3304425;
 RA Bodmer M.W., Angal S., Varranont G.T., Harris T.J.R., Lyons A.,
 RA King D.J., Pieroni G., Riviere C., Verger R., Lowe P.A.;
 RT "Molecular cloning of a human gastric lipase and expression of the
 RT enzyme in yeast.";

RL Biochim. Biophys. Acta 909:237-244(1987).
 RN [2]
 RP SEQUENCE OF 20-45.
 RX MEDLINE=89325292; PubMed=2753032;
 RA Bernbaeck S., Blaechberg L.;
 RT "Human gastric lipase. The N-terminal tetrapeptide is essential for
 RT lipid binding and lipase activity.";
 RL Eur. J. Biochem. 182:495-499(1989).
 CC -!- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O - diacylglycerol + a
 CC fatty acid anion.
 CC -!- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
 CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X05997; CAA29413.1; -
 CC EMBL; X05997; CAA29414.1; ALT_INIT.
 CC EMBL; A01046; CAA00125.1; -
 CC EMBL; A12714; CAA01053.1; -
 CC PIR; S04942; S04942.
 CC PIR; S07145; S07145.
 CC MIM; 601980; -
 CC InterPro; IPR000073; Abhydrolase.
 CC InterPro; IPR000379; Est_lip_thioest_actsite.
 CC InterPro; IPR000734; Lipase.
 CC Pfam; PF00561; abhydrolase.1.
 CC PROSITE; PS00120; Lipase_SER; 1.
 CC Hydrolyase; Lipid degradation; Signal; Glycoprotein; Polymorphism.
 KW SIGNAL
 FT CHAIN 1 19
 FT ACT_SITE 20 398 TRIACYLGLYCEROL LIPASE, GASTRIC.
 FT ACT_SITE 172 172 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CARBOHYD 34 34 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 99 99 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 271 271 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 327 327 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARIANT 161 161 T -> A (IN DBSNP:814628).
 FT SEQUENCE 398 AA; 45237 MW; CD3EE1621C014F0F CRC64;
 SQ
 Query Match 33.3%; Score 156; DB 1; Length 398;
 Best Local Similarity 47.0%; Pred. No. 1.7e-10;
 Matches 31; Conservative 13; Mismatches 20; Indels 2; Gaps 2;
 QY 10 YNNWSWDDLVLVNDLPAMVDVVKQTGQKP-HYVGHSMGTVALAAAFSEG-RVVSQKSA 67
 Db :||:||||: ||||| :||:||||: ||||| ||||| ||||| ||||| :||: |||||
 137 FWAFFDEMAKYDLPATIDFIVKKTGQKQLHYVGHSGQTTGTFIAFSTNPISAKRIKTFY 196
 QY 68 LLTPVA 73
 Db |||||
 197 ALAPVA 202
 RESULT 3
 LIPG_CANFA STANDARD; PRT; 398 AA.
 ID P80035; O02857;
 AC 01-NOV-1991 (Rel. 20, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Triacylglycerol lipase, gastric precursor (EC 3.1.1.3) (Gastric
 DE lipase) (GL).
 GN LIFF.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.


```

OX NCBI_TaxID=9615;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE-Stomach;
RA Vaganay S., Joliff G., Bertaux O., Toselli E., Devignes M.D.,
RA Benicourt C.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE OF 20-59.
RX MEDLINE=92037652; PubMed=1935982;
RA Carriere F., Moreau H., Raphael V., Laugier R., Benicourt C.,
RA Junien J.-L., Verger R.;
RT "Purification and biochemical characterization of dog gastric
RL lipase.";
RL Euf. J. Biochem. 202:75-83(1991).
CC -1- CATALYTIC ACTIVITY: Triacylglycerol + a
CC fatty acid anion.
CC -1- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y13899; CAA74198.1; -.
DR PIR; S19539; S19539.
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR000379; Est_lip_thioest_actsite.
DR Pfam; PF00561; abhydrolase.1.
DR PROSITE; PS00120; LIPASE_SER; 1.
DR Hydrolase; Lipid degradation; Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 1 19
FT ACT_SITE 172 172 TRIACYLGLYCEROL LIPASE, GASTRIC.
FT ACT_SITE 372 372 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 34 34 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 39 39 I -> T (IN REF. 2).
SQ SEQUENCE 398 AA; 45130 MW; E04D62F7518E386C CRC64;

Query Match 32.9%; Score 154; DB 1; Length 398;
Best Local Similarity 45.5%; Pred. No. 2.9e-10;
Matches 30; Conservative 15; Mismatches 19; Indels 2; Gaps 2;

Qy 10 YWNWSDDLVLVNDLPAMVDVFFVVKQTGQ-KPHYVGHSMGTALVALAAAFSEG-RVVSOLKSA 67
Db 137 FWAESFDEMAKYDLPATIDFILKKTGQDKLHYVGHSGQTGTFIAFSTNPKLAKIKTFY 196

Qy 68 LLTPVA 73
Db 197 ALAPVA 202

RESULT 4
LIPG_RAT STANDARD; PRT; 395 AA.
AC P04634;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Triacylglycerol lipase, lingual precursor (EC 3.1.1.3) (Lingual
DE lipase).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

```

```

OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=95215587; PubMed=3839077;
RA Docherty A.J.P., Bodmer M.W., Angal S., Verger R., Riviere C.,
RA Lowe P.A., Lyons A., Emtage J.S., Harris T.J.R.;
RT "Molecular cloning and nucleotide sequence of rat lingual lipase
RT cDNA.";
RL Nucleic Acids Res. 13:1891-1903(1985).
CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a
CC fatty acid anion.
CC -1- MISCELLANEOUS: LINGUAL LIPASE IS SECRETED BY THE SEROUS (VON
CC EBNER'S) GLANDS AT THE BACK OF THE RAT TONGUE.
CC -1- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02309; CAA26179.1; -.
DR EMBL; A01157; CAA00136.1; -.
DR PIR; A23045; LIPTT.
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR000379; Est_lip_thioest_actsite.
DR InterPro; IPR000734; Lipase.
DR Pfam; PF00561; abhydrolase.1.
DR PROSITE; PS00120; LIPASE_SER; 1.
DR Hydrolase; Lipid degradation; Glycoprotein; Signal.
FT SIGNAL 1 18
FT CHAIN 1 18
FT ACT_SITE 171 171 TRIACYLGLYCEROL LIPASE, LINGUAL.
FT ACT_SITE 371 371 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 33 33 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 395 AA; 44588 MW; E601854A923522EA CRC64;

Query Match 32.5%; Score 152; DB 1; Length 395;
Best Local Similarity 40.5%; Pred. No. 4.9e-10;
Matches 30; Conservative 17; Mismatches 25; Indels 2; Gaps 2;

Qy 10 YWNWSDDLVLVNDLPAMVDVFFVVKQTGQ-KPHYVGHSMGTALVALAAAFSEG-RVVSOLKSA 67
Db 136 FWAESFDEMAKYDLPATINFIQVKTQGEKIHYVGHSGQTGTFIAFSTNPTLAKIKTFY 195

Qy 68 LLTPVAYLXHNXP 81
Db 196 ALAPVATVKYQSP 209

RESULT 5
LIPG_BOVIN STANDARD; PRT; 397 AA.
ID LIPG_BOVIN
AC Q29458;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Triacylglycerol lipase, pregastric precursor (EC 3.1.1.3) (Pregastric
DE lipase) (GL) (Pregastric esterase) (pGE).
GN LIPF.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bos.
OX NCBI_TaxID=9913;
RN (1)

```



```
DR PIR: A39315; A39315.
DR MW: 278000; -.
DR InterPro: IPR000073; Abhydrolase.
DR InterPro: IPR000379; Est_lip_thioest_actsite.
DR InterPro: IPR000734; Lipase.
DR Pfam: PF00561; abhydrolase; 1.
DR PROSITE: PS00120; LIPASE_SER; 1.
KW Hydrolase; Lipid degradation; Glycoprotein; Signal; Lysosome;
KW Disease mutation; Polymorphism.
FT SIGNAL 1 21
FT CHAIN 22 399
FT ACT_SITE 174 174
FT ACT_SITE 374 374
FT CARBOHYD 36 36
FT CARBOHYD 72 72
FT CARBOHYD 101 101
FT CARBOHYD 161 161
FT CARBOHYD 273 273
FT CARBOHYD 321 321
FT VARIANT 16 16
FT VARIANT 129 129
FT VARIANT 129 129
FT VARIANT 129 129
FT VARIANT 200 200
FT CONFLICT 23 23
FT CONFLICT 29 29
FT SEQUENCE 399 AA; 45415 MW; 55F27391306B609A CRC64;

Query Match 29.18; Score 136; DB 1; Length 399;
Best Local Similarity 35.48; Pred. No. 3.4e-08;
Matches 28; Conservative 18; Mismatches 31; Indels 2; Gaps 2;

Qy 10 YNMSWDDLVDLPAMVDFVYVKTGQKP-HYVGHSMGTLVALAASF-GRVVSQKSA 67
Db 139 FWAFSDEMAKYDLPASINFLNKTGQEQVYVGHSGQTGTFIAFSQPELAKRMKF 198
Qy 68 LLTPVAYLXHNXP 86
Db 199 ALGPVAVAFCTSPMAKLG 217

RESULT 7
LICH_MOUSE STANDARD; PRT; 397 AA.
AC Q9Z0M5;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Lysosomal acid lipase/cholesteryl ester hydrolase precursor
DE (EC 3.1.1.13) (LAL) (Acid cholesteryl ester hydrolase) (Sterol
DE esterase) (Lipase A) (Cholesteryl esterase).
GN LIPA OR LIPL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
RA MEDLINE=96363957; Pubmed=8725147;
RX Du H., Witte D.P., Grabowski G.A.;
RT "Issue and cellular specific expression of murine lysosomal acid
RT lipase mRNA and protein."
RL J. Lipid Res. 37:937-949(1996).
CC -1- FUNCTION: CRUCIAL FOR THE INTRACELLULAR HYDROLYSIS OF CHOLESTERYL
CC ESTERS AND TRIGLYCERIDES THAT HAVE BEEN INTERNALIZED VIA RECEPTOR-
CC MEDIATED ENDOCYTOSIS OF LIPOPROTEIN PARTICLES. IMPORTANT IN
CC MEDIATING THE EFFECT OF LDL (LOW DENSITY LIPOPROTEIN) UPTAKE ON
```

```
CC SUPPRESSION OF HYDROXYMETHYLGUTARYL-COA REDUCTASE AND ACTIVATION
CC OF ENDOGENOUS CELLULAR CHOLESTERYL ESTER FORMATION (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: A sterol ester + H(2)O -> a sterol + a fatty
CC acid.
CC -1- SUBCELLULAR LOCATION: Lysosomal.
CC -1- TISSUE SPECIFICITY: Expressed at low levels in most tissues. High
CC level expression is found in hepatocytes and splenic and thymic
CC cells. Very high level expression is observed in cells of the
CC small intestinal villi, the zona fasciculata and reticularis of
CC the adrenal cortex, pancreatic acini, and renal tubular
CC epithelium.
CC -1- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z31689; CAA83494.1; -.
CC MGD; MG1;96789; Lip1.
CC InterPro: IPR000073; Abhydrolase.
CC InterPro: IPR000379; Est_lip_thioest_actsite.
CC InterPro: IPR000734; Lipase.
CC Pfam: PF00561; abhydrolase; 1.
CC PROSITE: PS00120; LIPASE_SER; 1.
KW Hydrolase; Lipid degradation; Glycoprotein; Signal; Lysosome.
FT SIGNAL 1 19
FT CHAIN 20 397
FT ACT_SITE 172 172
FT ACT_SITE 372 372
FT CARBOHYD 34 34
FT CARBOHYD 99 99
FT CARBOHYD 159 159
FT CARBOHYD 271 271
FT CARBOHYD 319 319
FT SEQUENCE 397 AA; 45551 MW; F886C39E1CCFA91F CRC64;

Query Match 27.6%; Score 129; DB 1; Length 397;
Best Local Similarity 33.8%; Pred. No. 2.1e-07;
Matches 25; Conservative 20; Mismatches 27; Indels 2; Gaps 2;

Qy 10 YNMSWDDLVDLPAMVDFVYVKTGQKP-HYVGHSMGTLVALAASF-GRVVSQKSA 67
Db 137 FWAFSDEMAKYDLPASINFLNKTGQEQVYVGHSGQTGTFIAFSQPELAKRMFL 196
Qy 68 LLTPVAYLXHNXP 81
Db 197 VLAPVLSLNFASGP 210

RESULT 8
LICH_RAT STANDARD; PRT; 397 AA.
AC Q64194;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Lysosomal acid lipase/cholesteryl ester hydrolase precursor
DE (EC 3.1.1.13) (LAL) (Acid cholesteryl ester hydrolase) (Sterol
DE esterase) (Lipase A) (Cholesteryl esterase).
GN LIPA OR LAL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
```


QY 56 EGR 58
|
Db 162 PDR 164

Search completed: October 10, 2002, 02:14:31
Job time: 214 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:10:32 ; Search time 14.1 Seconds
(without alignments)
613.336 Million cell updates/sec

Title: US-09-699-652-12
Perfect score: 468
Sequence: 1 SSCTLRFLYWNWSDDLTV.....PVAYLHXNXPXNPGWPXR 90

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	170	36.3	344	2 E84526	probable lysosomal
2	156	33.3	398	2 S07145	triacylglycerol li
3	152	32.5	395	1 L1RTT	triacylglycerol li
4	147.5	31.5	411	2 T22290	hypothetical prote
5	145.5	31.1	405	2 H88930	protein R1G11.14
6	145.5	31.1	405	2 T22675	hypothetical prote
7	145	31.0	397	1 JG4017	triacylglycerol li
8	139.5	29.8	403	2 T33198	hypothetical prote
9	136	29.1	399	2 S41408	lysosomal acid lip
10	136	29.1	399	2 G01416	lysosomal acid lip
11	126	26.9	559	2 J70949	egg-specific prote
12	123	26.3	426	2 T20480	hypothetical prote
13	122	26.1	411	2 G89074	protein R0A48.5 li
14	118.5	25.3	443	2 T39540	triglyceride lipas
15	101	21.6	467	2 T41053	triglyceride lipas
16	95.5	20.4	413	2 T43170	probable triacylg
17	93.5	20.0	336	2 F83425	probable esterase/
18	92.5	19.8	460	2 G96764	unknown protein F2
19	90.5	19.3	538	2 S64842	probable membrane
20	86.5	18.5	523	2 D86284	F9L1.1 protein - A
21	85	18.2	207	2 C97494	hypothetical prote
22	85	18.2	207	2 AE2712	conserved hypothe
23	79.5	17.0	1585	2 T31611	hypothetical prote
24	77	16.5	369	2 E75620	hypothetical prote
25	76	16.2	438	2 G87675	arylesterase-relat
26	75	16.0	241	2 C84612	2-hydroxy-6-oxonep
27	75	16.0	241	2 B71903	hypothetical prote
28	74.5	15.9	460	2 T39443	probable triglycer
29	73	15.6	353	2 B90424	esterase, probable

30	73	15.6	605	2 I39837	dnak-type molecula
31	71.5	15.3	374	1 D42462	dihydrolipoamide S
32	69	14.7	548	2 S37969	probable triacylg
33	68	14.5	236	2 F75443	probable hydrolase
34	68	14.5	573	2 S64754	probable membrane
35	67.5	14.4	204	2 S55916	ribosomal protein
36	67.5	14.4	220	2 F84790	40S ribosomal prot
37	65	13.9	220	2 S32111	MPT51 protein - My
38	65	13.9	284	2 T00809	probable esterase
39	65	13.9	292	2 T33322	hypothetical prote
40	65	13.9	301	2 B86921	antigen 85C, mycol
41	64.5	13.8	1097	2 T40678	hypothetical prote
42	64	13.7	204	1 R3RT5	ribosomal protein
43	64	13.7	290	2 A99516	esterase/lipase 2
44	63.5	13.6	133	2 S12097	rfai protein - Sal
45	63.5	13.6	193	2 D71425	probable ATSEH - A

ALIGNMENTS

RESULT 1

E84526
probable lysosomal acid lipase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: E84526
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
M.; Koo, H.; Moffat, R.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: E84526
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-344 <STO>
A:Cross-references: GB:A802093; NID:94585908; PIDN:AAD25569.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g15230
A:Map position: 2

Query Match 36.3%; Score 170; DB 2; Length 344;
Best Local Similarity 38.9%; Pred. No. 1.2e-11;
Matches 28; Conservative 17; Mismatches 27; Indels 0; Gaps 0;

QY 10 YNWSWDDLVMNDLPAMVDVVYKQGPQPHYVGHSMGTLLVALAASFSEGRVVSQKSAALL 69

Db 132 FWDWSWQDLAMYDLAEIMQIYLSISNSKIFLVGHSGQTIMSPAALTQPHVAEWVEAALL 191

QY 70 TPVAYLXHXNXP 81

Db 192 CPISYLDHVTAP 203

RESULT 2

S07145
triacylglycerol lipase (EC 3.1.1.3) precursor, gastric - human

C:Species: Homo sapiens (man)

C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 18-Jun-1999

C:Accession: S07145; S27102; S04942

R:Bohmer, M.W.; Angal, S.; Varranton, G.T.; Harris, T.J.R.; Lyons, A.; King, D.J.; P

Biochim. Biophys. Acta 909, 237-244, 1987

A:Title: Molecular cloning of a human gastric lipase and expression of the enzyme in

A:Reference number: S07145; MUID:87299724

A:Accession: S07145

A:Molecule type: mRNA

A:Residues: 1-398 <B0D1>

A:Cross-references: EMBL:X05997; NID:g31771; PIDN:CAA29413.1; PID:g758063

A:Accession: S27102

A:Molecule type: protein

A:Residues: 20-24,'X',26,'X',28-33,'X',35-42,'X',44,'X',46-48,'X',50,'X',52,'X',54-56

submitted to the EMBL Data Library, April 1996

A; Reference number: 219598

C:Superfamily: triacylglycerol lipase, lingual
C:Keywords: carboxylic ester hydrolase; glycoprotein

Query Match 29.1%; Score 136; DB 2; Length 399;
Best Local Similarity 35.4%; Pred. No. 1e-07; 31; Indels 2; Gaps 2;
Matches 28; Conservative 18; Mismatches 31;
QY 10 YNWSMDDLAVNDLPAMVDFVVKQTGQK-PHYVGHSMGTLVALAAAFSE-GRVVSOLKSA 67
DB 139 FNAFSVDENAKYDLPAISINFILNKYGQEQVYVGHSGQTTFGIAFSQIPELAKRIKMF 198
QY 68 LLTPVAYLXHXNPNKPG 86
DB 199 ALGPVASVAFCTSPMAKLG 217

RESULT 10
G01416
lysosomal acid lipase - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 18-Jun-1999
C:Accession: G01416

R:Du, H.
submitted to the EMBL Data Library, April 1994
A:Reference number: G06919
A:Accession: G01416
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-399 <DUX>
A:Cross-references: EMBL:008464; NID:g505052; PIDN:AAB60328.1; PID:g505053
C:Superfamily: triacylglycerol lipase, lingual
C:Keywords: glycoprotein

Query Match 29.1%; Score 136; DB 2; Length 399;
Best Local Similarity 35.4%; Pred. No. 1e-07; 31; Indels 2; Gaps 2;
Matches 28; Conservative 18; Mismatches 31;
QY 10 YNWSMDDLAVNDLPAMVDFVVKQTGQK-PHYVGHSMGTLVALAAAFSE-GRVVSOLKSA 67
DB 139 FNAFSVDENAKYDLPAISINFILNKYGQEQVYVGHSGQTTFGIAFSQIPELAKRIKMF 198
QY 68 LLTPVAYLXHXNPNKPG 86
DB 199 ALGPVASVAFCTSPMAKLG 217

RESULT 11
JT0949
egg-specific protein - silkworm
C:Species: Bombyx mori (silkworm)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 26-Feb-1998
C:Accession: JT0949; A28527

R:Sato, Y.; Yamashita, O.
submitted to JIPID, September 1991
A:Reference number: JT0949
A:Accession: JT0949
A:Molecule type: DNA
A:Residues: 1-559 <SAT>
A:Experimental source: larva
A:Note: this protein is a homotrimer
R:Indrasith, L.S.; Sasaki, T.; Yamashita, O.
J. Biol. Chem. 263, 1045-1051, 1988
A:Title: A unique protease responsible for selective degradation of a yolk protein in Bombyx mori
A:Reference number: A28527; MUID:88087166
A:Accession: A28527
A:Molecule type: protein
A:Residues: 19-40;133-144, 'T',146-152;229-248 <IND>
C:Genetics:
A:Map position: 19
C:Keywords: egg yolk; homotrimer

F:132-133/Cleavage site: Lys-Asn (egg 30.5K cysteine proteinase) #status experimental

F:228-229/Cleavage site: Arg-Asp (egg 30.5K cysteine proteinase) #status experimental

Query Match 26.9%; Score 126; DB 2; Length 559;
Best Local Similarity 33.8%; Pred. No. 2.1e-06;
Matches 25; Conservative 19; Mismatches 28; Indels 2; Gaps 2;
QY 10 YNWSMDDLAVNDLPAMVDFVVKQTGQ-PHYVGHSMGTLVALAAAFSEGRVVSQ-LKSA 67
DB 293 FWKFSNDEALHDLPAIDHVLDISQERLHVIGHSOGATTFALMSEQPSYNEKIVSMH 352
QY 68 LLTPVAYLXHXNXP 81
DB 353 ALSPIVYMNVRSP 366

RESULT 12
T20480
hypothetical protein F01G10.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T20480
R:Hembry, C.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19280
A:Accession: T20480
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-426 <WIL>
A:Cross-references: EMBL:281055; PIDN:CAB02896.1; GSPDB:GN000022; CESP:F01G10.7
A:Experimental source: clone F01G10
C:Genetics:
A:Gene: CESP:F01G10.7
A:Map position: 4
A:Introns: 29/3; 64/1; 186/3; 347/2
C:Superfamily: triacylglycerol lipase, lingual

Query Match 26.3%; Score 123; DB 2; Length 426;
Best Local Similarity 34.7%; Pred. No. 3.4e-06;
Matches 25; Conservative 17; Mismatches 28; Indels 2; Gaps 2;
QY 8 QLYWNWSDDLAVNDLPAMVDFVVKQTGQK-PHYVGHSMGTLVALAAAFSEG-RVVSOLKS 65
DB 130 QKFNFNTWQOMSEFDLTASVDLVKTKQEFLYLGHSGQTMIMFSRLAENPEFAKKIRH 189
QY 66 AALLTPVAYLXH 77
DB 190 FHALAPVATVSH 201

RESULT 13
G89074
protein K04A8.5 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C:Accession: G89074
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_-
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A:Accession: G89074
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-411 <STO>
A:Cross-references: GB:chr_V; PIDN:AAC48051.1; PID:gl658354; GSPDB:GN000023; CESP:K04A
C:Genetics:
A:Gene: K04A8.5
A:Map position: 5
C:Superfamily: triacylglycerol lipase, lingual

Query Match 26.1%; Score 122; DB 2; Length 411;
 Best Local Similarity 33.3%; Pred. No. 4.3e-06;
 Matches 24; Conservative 17; Mismatches 29; Indels 2; Gaps 2;
 QY 10 YNWSDDLVDLPAMVDFVVKQTGP-HVYGHSMGTLVALAAFS-EGRVVSOLKSAA 67
 DB 142 FWDMSWDQISEYDLPAMIGKALEISQESLYTGTSLTMTFAKLSTDPKFSRKIKKYF 201
 QY 68 LPTPVAYLXHXN 79
 DB 202 ALAPIGSINHAH 213

RESULT 14
 T39540
 triglyceride lipase-cholesterol esterase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C:Accession: T39540
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.
 submitted to the EMBL Data Library, February 1998
 A:Reference number: Z21862
 A:Accession: T39540
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-443 <WOO>
 A:Cross-references: EMBL:AL021748; PIDN:CAA16863.1; GSPDB:GN00067; SPDB:SPBC16A3.12c
 A:Experimental source: strain 972h-; cosmid cl6A3
 C:Genetics:
 A:Gene: SPDB:SPBC16A3.12c
 A:Map position: 2
 A:Introns: 11/2; 152/3; 282/1; 398/3
 C:Superfamily: triacylglycerol lipase, lingual

Query Match 25.3%; Score 118.5; DB 2; Length 443;
 Best Local Similarity 44.7%; Pred. No. 1.2e-05;
 Matches 21; Conservative 14; Mismatches 11; Indels 1; Gaps 1;
 QY 10 YNWSDDLVDLPAMVDFVVKQTG-QKPHYVGHSMGTLVALAAFS 55
 DB 175 FWNFLDDMFIDPTVDYILRETGREKLNIFYGSGTAQAQAALS 221

RESULT 15
 T41053
 triglyceride lipase-cholesterol esterase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T41053
 R:Hilbert, H.; Duasterhoeft, A.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, August 1998
 A:Reference number: Z21967
 A:Accession: T41053
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-467 <HIL>
 A:Cross-references: EMBL:AL031324; PIDN:CAA20447.1; GSPDB:GN00068; SPDB:SPCC1672.09
 A:Experimental source: strain 972h-; cosmid cl672
 C:Genetics:
 A:Gene: SPDB:SPCC1672.09
 A:Map position: 3
 A:Introns: 12/2; 164/3; 294/1

Query Match 21.6%; Score 101; DB 2; Length 467;
 Best Local Similarity 32.5%; Pred. No. 0.0012;
 Matches 25; Conservative 17; Mismatches 25; Indels 10; Gaps 3;
 QY 5 LRF----QLYNNWSDDLVDLPAMVDFVVKQTG-QKPHYVGHSMGTLVALAAFS----- 55
 DB 178 LRFSDTKREFWDFSIDFAQYDIPDTIDYILKTSQTKLTYIGFSQGTAAQAFASLSIHPL 237

QY 56 -EGRVVSOLKSAAALITP 71
 DB 238 LNDKINSILIALAPAISP 254

Search completed: October 10, 2002, 02:14:05
 Job time: 213 sec

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	279	59.6	410	6	US-10-219-999-43231	Sequence 43321, A	
2	233	49.8	266	6	US-10-219-999-42658	Sequence 42658, A	
3	233	49.8	399	6	US-10-219-999-35161	Sequence 35161, A	
4	181	38.7	393	6	US-10-219-999-54283	Sequence 54283, A	
5	172	36.8	394	5	US-09-791-537-94868	Sequence 94868, A	
6	170	36.3	127	6	US-10-056-744B-5	Sequence 5, Appl	
7	170	36.3	344	5	US-09-791-537-49394	Sequence 49394, A	
8	156	33.3	371	5	US-09-791-537-50507	Sequence 50507, A	
9	156	33.3	392	5	US-09-791-537-123314	Sequence 123314, A	
10	156	33.3	398	5	US-09-791-537-93233	Sequence 93233, A	
11	156	33.3	398	5	US-09-759-130B-445	Sequence 445, Appl	
12	156	33.3	398	6	US-10-002-945-66	Sequence 66, Appl	
13	155	33.1	182	6	US-10-219-999-52422	Sequence 52422, A	
14	154	32.9	398	5	US-09-791-537-112004	Sequence 112004, A	
15	152	32.5	395	5	US-09-791-537-122784	Sequence 122784, A	
16	149	31.8	398	5	US-09-791-537-47839	Sequence 47839, A	
17	147.5	31.5	411	5	US-09-791-537-37594	Sequence 37594, A	
18	145.5	31.1	405	5	US-09-791-537-36768	Sequence 36768, A	
19	145.5	31.1	405	5	US-09-791-537-55922	Sequence 55922, A	
20	145	31.0	397	5	US-09-791-537-126910	Sequence 126910, A	
21	145	31.0	399	6	US-10-221-097-35	Sequence 35, Appl	
22	141	30.1	399	5	US-09-791-537-110191	Sequence 110191, A	
23	141	30.1	435	5	US-09-791-537-33752	Sequence 33752, A	
24	141	30.1	616	5	US-09-791-537-9393	Sequence 9393, Appl	
25	139.5	29.8	403	5	US-09-791-537-89046	Sequence 89046, A	
26	138	29.5	221	5	US-09-759-130B-420	Sequence 420, Appl	

```

; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)C
; CURRENT APPLICATION NUMBER: US/10/219,999
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/324,109
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/312,544
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 63520
; SEQ ID NO 42658
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Glycine max
US-10-219-999-42658

```

Query Match 49.88; Score 233; DB 6; Length 266;
Best Local Similarity 55.68; Pred. No. 1.9e-21;
Matches 40; Conservative 12; Mismatches 20; Indels 0; Gaps 0;

Qy	10	YNNWSUDDLVLVNDLPAWVDFVVKOTGOKPHVGHUSNGTLVALAASFEGSEGRVVSQLSKAALL	69
Db	3	YWDMSDELARYDLPAFVQYVYINOTGORMHYAGHSGLTLMVLADLSRGKLLDMLRSKAALL	62
Qy	70	TPVAYLXHXNXP	81
Db	63	CPIAHLNHVTSP	74

```

RESULT      3
US-10-219-999-35161
; Sequence 35161, Application US/10219999
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)C
; CURRENT APPLICATION NUMBER: US/10/219,999
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/324,109
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/312,544
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 63520
; SEQ ID NO 35161
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Glycine max
US-10-219-999-35161

```

Query Match 49.8%; Score 233; DB 6; Length 399;
Best Local Similarity 55.6%; Pred. No. 3.1e-21;
Matches 40; Conservative 12; Mismatches 20; Indels 0; Gaps 0;

[illegible]

RESULT 4
US-10-219-999-54283
; Sequence 54283, Application US/10219999
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei

```

: APPLICANT: Edgerton, Michael D
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Kovalic, David K.
: APPLICANT: Liu, Jingdong
: APPLICANT: Stein, Joshua
: TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
: FILE REFERENCE: 38-10(52726)C
: CURRENT APPLICATION NUMBER: US/10/219,999
: CURRENT FILING DATE: 2002-08-15
: PRIOR APPLICATION NUMBER: US 60/324,109
: PRIOR FILING DATE: 2001-09-21
: PRIOR APPLICATION NUMBER: US 60/312,544
: PRIOR FILING DATE: 2001-08-15
: NUMBER OF SEQ ID NOS: 63520
: SEQ ID NO 54283

```

Query Match	38.7%	Score	181;	DB	6;	Length	393;
Best Local Similarity	44.2%;	Pred. No.	1.4e-14;				
Matches	34;	Conservative	15;	Mismatches	28;	Indels	0;

[illegible]

```

RESULT      5
US-09-791-537-94868
; Sequence 94868, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 94868
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-791-537-94868

```

Query Match	36.8%	Score 172;	DB 5;
Best Local Similarity	44.2%	Pred. No. 2.1e-13;	Length 394;
Matches 34;	Conservative	18;	Mismatches 23;
			Indels 2;
			Gaps 2;

[illegible]

RESULT 6
US-10-056-744B-5
; Sequence 5, Application US/10056744B
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana

QY 68 LLTPVA 73

THIS PAGE BLANK (USPTO)


```

; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-DI-66-TMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-073-674-3

Query Match 32.9%; Score 154; DB 2; Length 379;
Best Local Similarity 45.5%; Pred. No. 4.4e-13;
Matches 30; Conservative 15; Mismatches 19; Indels 2; Gaps

QY 10 YNWSDDLLVNDLPAMDFVVKQTQG-RPHYYGVHSMGTLVALAATSEG-RVVSQLKSA 67
Db 118 FWAFSDEMAKYDLPATIDFILAKTGDKLHYVHSQGTIGTIAFTNPKLAKRIKTFY 177

QY 68 LLTPVA 73
Db 178 ALAPVA 183

RESULT 5
US-08-227-108-5
; Sequence 5, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-227-108-5

```



```

; CITY: Zurich
; STATE: none
; COUNTRY: Switzerland
; ZIP: (note: this is an international post code) CH-8092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/857,224B
; FILING DATE: 03/25/92
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA: none
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (International) 41 1 632 2830
; TELEFAX: (International) 41 1 262 2437
; TELEX: none
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 344
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
; ORIGINAL SOURCE:
; ORGANISM: Scizosaccharomyces pombe
; FEATURE: Alcohol dehydrogenase, Table 3 Column 15
; PUBLICATION INFORMATION:
; AUTHORS: Joernvall, H.
; AUTHORS: Persson, M.
; AUTHORS: Jeffery, J.
; TITLE: Alcohol dehydrogenases
; JOURNAL: Proceedings of the National Academy of Sciences, USA
; VOLUME: 78
; PAGES: 4226-4230
; DATE: 1981
; US-07-857-224B-92

Query Match 13.0%; Score 61; DB 2; Length 344;
Best Local Similarity 35.9%; Pred. No. 2.7;
Matches 23; Conservative 6; Mismatches 27; Indels 8; Gaps

QY 19 VVNDLPA-----MVDEVVVKQTGQKPHVVGSHMGTLVALAAFSEGRVVSOLKSAALLT- 70
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 265 VVGMFAGAKCCSDVFNQVVKSIKGSYVGNRADTREALDFARGLVKSPIKVVGLSTL 324
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

QY 71 PVAY 74
   ||::||
Db 325 PEIY 328

RESULT 15
US-08-095-734-2
; Sequence 2, Application US/08095734
; Patent No. 5807723
; GENERAL INFORMATION:
; APPLICANT: Aldovini, Anna
; APPLICANT: Young, Richard A.
; TITLE OF INVENTION: Homologously Recombinant Slow Growing
; TITLE OF INVENTION: Mycobacteria and Uses Therefor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

Search completed: October 10, 2002, 02:10:54
Job time: 662 sec

THIS PAGE BLANK (USPTO)